Univerzita Karlova
Přírodovědecká fakulta

Studijní program: Biochemie
Studijní obor: Biochemie

Iva Hánová

Nové inhibiční mechanismy regulace aktivity katepsinu D
New inhibition mechanisms of regulation of cathepsin D activity

Disertační práce

Školitel: RNDr. Michael Mareš, CSc.
Praha 2020

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, dne

……

Podpis

Poděkování

Ráda bych zde poděkovala všem, kteří mi pomáhali během studia.

Děkuji školiteli Michaelu Marešovi za vedení celé práce, ochotu a čas, který mi věnoval. Dále bych ráda poděkovala Martinu Hornovi a Martinu Mášovi za praktické rady a velkou trpělivost. Jiřímu Bryndovi bych ráda poděkovala za řešení struktur, Zdeňkovi Voburkovi za N-koncové sekvenování proteinů, Mirce Blechové za syntézu peptidů a Janě Starkové za pomoc s tkáňovými kulturami. Členům skupiny Hmotnostní spektrometrie děkuji za určení proteinových sekvencí. Celé skupině Katepsinové proteasy v patologii děkuji za všestrannou pomoc a příjemné pracovní prostředí.

V neposlední řadě bych ráda poděkovala mým nejmilovanějším za podporu a trpělivost během dlouhého studia.

Obsah

Seznam zkratek……..……...……..…..1
Abstrakt (ČJ)……..……..……..2
Abstract (EN)……...…….……3
1 Teoretický úvod.....4
1.1 Aspartátové proteasy a katepsin D.....4
1.1.1 Klasifikace proteas.....4
1.1.2 Aspartátové proteasy.....5
1.1.2.1 Rozšíření a role aspartátových proteas rodiny pepsinu.....5
1.1.2.2 Struktura aspartátových proteas rodiny pepsinu.....6
1.1.2.3 Katalytický mechanismus aspartátových proteas.....7
1.1.2.4 Substrátová specifita aspartátových proteas rodiny pepsinu.....8
1.1.2.5 Regulace aktivity aspartátových proteas rodiny pepsinu.....9
1.1.3 Proteasy typu katepsin D.....12
1.1.3.1 Funkce proteas typu katepsin D.....12
1.2 Lidský katepsin D.....14
1.2.1 Fyziologie lidského katepsinu D.....14
1.2.1.1 Lokalizace a fyziologické funkce lidského katepsinu D.....14
1.2.1.2 Syntéza a aktivace lidského katepsinu D.....15
1.2.2 Patofyziologie lidského katepsinu D.....18
1.2.2.1 Katepsin D v nádorových onemocněních.....18
1.2.2.2 Neurodegenerativní onemocnění spojovaná s katepsinem D.....20
1.2.2.3 Role katepsinu D v ateroskleróze.....23
1.2.3 Biochemická charakterizace lidského katepsinu D.....24
1.2.3.1 Struktura lidského katepsinu D.....24
1.2.3.2 Proteolytická aktivita lidského katepsinu D.....25
1.3 Katepsin D z parazitických klíšťat.....26
1.3.1 Taxonomie, morfologie a životní cyklus klíštěte obecného.....26
1.3.1.1 Taxonomie a morfologie klíštěte obecného.....26
1.3.1.2 Životní cyklus klíšťat.....26
1.3.2 Trávicí systém a trávicí proteasy klíšťat.....27
1.3.2.1 Trávicí systém klíštěte.....27
1.3.2.2 Trávicí proteasy klíštěte.....29
1.3.3 Katepsin D1 z klíštěte obecného *Ixodes ricinus*.....30
1.3.3.1 Fyziologie katepsinu D1 z klíštěte obecného.....31
1.3.3.2 Biochemická charakterizace katepsinu D1 z klíštěte obecného.....31
2 Cíle disertační práce.....32
3 Materiály a metody.....33
3.1 Materiály a laboratorní vybavení.....33
3.2 Metody.....33
4 Výsledky.....35

4.1 Publikace č.1: Complex modulation of peptidolytic activity of cathepsin D by sphingolipids 36
4.1.1 Souhrn 36
4.1.2 Publikace č. 1 38
4.2 Publikace č. 2: Novel structural mechanism of allosteric regulation of aspartic peptidases via an evolutionarily conserved exosite 50
4.2.1 Souhrn 50
4.2.2 Publikace č. 2 52
4.3 Publikace č. 3: Biomimetic macrocyclic inhibitors of human cathepsin D: structure-activity relationship and binding mode analysis 81
4.3.1 Souhrn 81
4.3.2 Publikace č. 3 83
5 Diskuze 111
6 Závěry 115
6.1 Prohlášení spoluautorů 117
7 Seznam použité literatury 118

Seznam zkratek

Å	Ångström, 10^{-10}m
APP	amyloidní prekurzorový protein
β-AP	β-amyloidní peptid
BACE1	β-sekretasa 1
ECM	extracelulární matrix
HIV	virus lidské imunodeficience
HPLC	„High Performance Liquid Chromatography"
IC_{50}	koncentrace inhibitoru potřebná k dosažení 50% inhibice aktivity enzymu
IrKatD1-3	isoformy 1-3 katepsinu D z klíštěte obecného (*Ixodes ricinus*)
KatD	katepsin D
k_{cat}	číslo přeměny
k_{cat}/K_M	katalytická účinnost enzymu
kDa	kilodalton
K_i	inhibiční konstanta
K_M	Michaelisova konstanta
LRP1	„Low-density lipoprotein Receptor-related Protein-1"
M6F-R	mannosa-6-fosfátový receptor
PDB	„Protein Data Bank", databáze proteinových struktur
PDI	„Potato Cathepsin D Inhibitor", inhibitor katepsinu D z brambor
PI3	„Pepsin-Inhibitor 3", inhibitor pepsinu ze škrkavky *(Ascaris suum*)
proKatD	prokatepsin D
proIrKatD1	proenzym katepsinu D1 z klíštěte obecného *(Ixodes ricinus)*
pseudoKatD	pseudokatepsin D
RP-HPLC	„Reverse Phase High Performace Liquid Chromatography"
SDS-PAGE	elektroforéza na polyakrylamidovém gelu v přítomnosti dodecylsíranu sodného

Nestandartní aminokyselinové zbytky:

NPh	p-nitrofenylalanin
Sta	statin

Abstrakt (ČJ)

Aspartátová proteasa katepsin D (KatD) je spojená s řadou patologií, a proto jsou zkoumány molekulární mechanismy regulace její aktivity, které mají potenciální využití v biomedicíně. Tato disertační práce je zaměřena na nové přirozené endogenní inhibitory KatD, analýzu jejich interakce a na odvození syntetických inhibičních biomimetik.

Byly identifikovány dvě skupiny inhibitorů KatD jako prvních specifických endogenních regulátorů tohoto enzymu. (1) Sfingolipidy jsou komplexními modulátory lidského KatD v závislosti na struktuře. Zatímco sfingosiny a ceramidy jsou inhibitory KatD, jejich fosforylované deriváty naopak fungují jako aktivátory KatD. Byla nalezena korelace mezi působením těchto sfingolipidů na KatD a jejich modulačním účinkem na nádorové buňky. (2) Analýzou KatD parazitárního původu byl identifikován nový mechanismus inhibice, který je konzervovaný u aspartátových proteas rodiny pepsinu. Autokatalytickým působením je z proenzymu u KatD uvolněn peptidový fragment, který působí jako alosterický inhibitor, jenž se váže do exomísta na povrchu aktivního enzymu. Dále byly připraveny syntetické makrocyklické inhibitory KatD, které mimikují vazebnou konformaci bakteriálního inhibitoru pepstatinu v aktivním místě aspartátových proteas. U těchto biomimetických inhibitorů lze na základě provedené analýzy strukturně-funkčních vztahů cíleně modifikovat jejich vlastnosti.

Disertační práce přináší významné informace o třech nových typech inhibitorů KatD a regulačních mechanismech, které mohou být využity pro racionální navrhování inhibičních léčiv u patologií spojených s aspartátovými proteasami typu KatD.

Abstract (EN)

The aspartic protease cathepsin D (CatD) is associated with numerous pathologies, and therefore the molecular mechanisms of its activation are studied for their potential uses in biomedicine. This dissertation thesis is focused on new, natural endogenous inhibitors of CatD, the analysis of their interaction, and the development of synthetic inhibitory biomimetics.

Two groups of inhibitors of CatD, which are the first specific endogenous regulators of this enzyme, have been identified. (1) Sphingolipids are complex modulators of human CatD, depending on their structure. While sphingosines and ceramides are inhibitors of CatD, their phosphorylated derivates act as activators of CatD. A correlation was found between the action of these sphingolipids on CatD and their modulatory effect on cancer cells. (2) Using the analysis of a CatD of parasitic origin, a new mechanism of inhibition was identified, which is conserved in aspartic proteases of the pepsin family. A peptide fragment is released autocatalytically from the zymogen of CatD, which then acts as an allosteric inhibitor, binding to an ”exosite” on the surface of the catalytically active enzyme. Furthermore, synthetic macrocyclic inhibitors of CatD were prepared, which mimic the binding conformation of the bacterial inhibitor pepstatin in the active site of aspartic proteases. Based on these results of the structure-function analysis, it is possible to purposefully modify the properties of these biomimetic inhibitors.

This thesis contains important new information about three new types of inhibitors of CatD and regulation mechanisms, which can be used for an intelligent design of inhibitory drugs for pathologies associated with CatD-like aspartic proteases.

1 Teoretický úvod

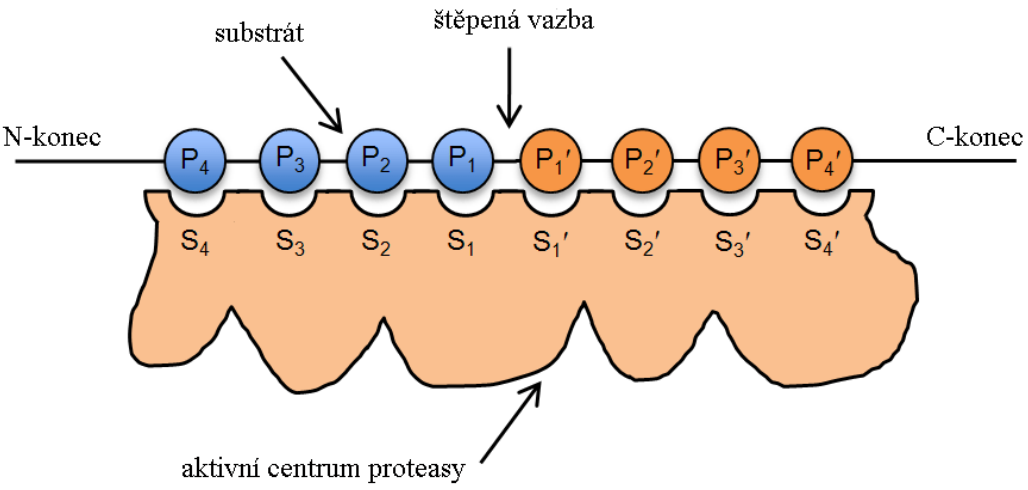
1.1 Aspartátové proteasy a katepsin D

Proteasy (nebo peptidasy) jsou enzymy, které hydrolyzují peptidovou vazbu proteinových a peptidových substrátů. Endopeptidasy štěpí peptidovou vazbu uvnitř (poly)peptidového řetězce substrátu. Naopak exopeptidasy odštěpí jednu aminokyselinu, dipeptid či tripeptid z konce substrátu. Exopeptidasy lze dále dělit na aminopeptidasy, které štěpí substrát z jeho N-konce, a karboxypeptidasy, které štěpí z C-konce substrátu. Podle počtu odštěpených aminokyselin se proteasy dělí na monopeptidasy, které odštěpují jednu aminokyselinu, dipeptidasy, které odštěpují dvě aminokyseliny atd[1,2].

1.1.1 Klasifikace proteas

Proteasy lze klasifikovat podle katalytického mechanismu a uspořádání klíčových aminokyselinových zbytků katalytického centra do sedmi tříd: aspartátové, glutamátové, serinové, cysteinové, threoninové, asparaginové a metaloproteasy[3]. U aspartátových, glutamátových a metaloproteas zahajuje štěpení peptidové vazby aktivovaná molekula vody. U serinových, cysteinových a threoninových proteas zahajují štěpení peptidové vazby postranní řetězce katalytického aminokyselinového zbytku v aktivním centru[2]. Na rozdíl od ostatních tříd proteas nejsou asparaginové proteasy hydrolasy ale lyasy, u kterých dochází k rozštěpení peptidové vazby eliminační reakcí[4].

Pro popis vazby (poly)peptidového substrátu do aktivního centra enzymu se využívá konvenční terminologie, kdy aminokyseliny substrátu, mezi kterými proteasa štěpí, jsou označovány jako pozice P_1 a P_1‘[5]. Dále jsou aminokyseliny označovány jako P_2, P_3 atd. od štěpené vazby směrem k N-konci substrátu a analogicky P_2‘, P_3‘ směrem od štěpené vazby k C-konci substrátu. Odpovídající vazebná podmísta v aktivním centru enzymu jsou označována jako S_1, S_2, S_3 atd., respektive S_1‘, S_2‘, S_3‘ atd[5]. Schematické znázornění aktivního centra a vazby substrátu je na obr. 1 na str. 5.



Obr. 1 Schematické znázornění vazby substrátu do aktivního centra proteasy. Zobrazeny jsou pozice jednotlivých aminokyselin substrátu (P_1-P_4 směrem od štěpené vazby k N-konci, P_1'-P_4'směrem od štěpené vazby k C-konci) a odpovídající podmísta v aktivním centru proteasy (S_1-S_4, respektive S_1'-S_4'). Substrát je rozštěpen mezi aminokyselinami P_1 a P_1'. Převzato a upraveno podle[5].

1.1.2 Aspartátové proteasy

Aspartátové proteasy jsou relativně málo zastoupenou třídou, řadí se do ní pouze 7% známých proteas[6]. Nicméně aspartátové proteasy hrají významnou roli v řadě patologických procesů, proto jsou podrobně studovány a jsou považovány za perspektivní cílové molekuly pro vývoj nových léčiv[7].

Aspartátové proteasy lze podle databáze MEROPS rozdělit do pěti klanů a dále třinácti rodin[3,8]. Proteasy jsou klasifikovány do rodin na základě homologie aminokyselinové sekvence. Proteasy v jednom klanu vykazují evoluční příbuznost podobnou terciární strukturu a také sekvenčním motivem v blízkosti katalytických zbytků. Nejvýznamnějším klanem aspartátových proteas je klan AA s rodinou A1 pepsinu a rodinou A2 retrovirálních proteas. Dále lze uvést klan AD s rodinou A22 proteas vázaných na buněčnou membránu obsahující např. proteasu presenilin. Tato práce se zaměřuje na proteolytické enzymy z rodiny pepsinu.

1.1.2.1 Rozšíření a role aspartátových proteas rodiny pepsinu

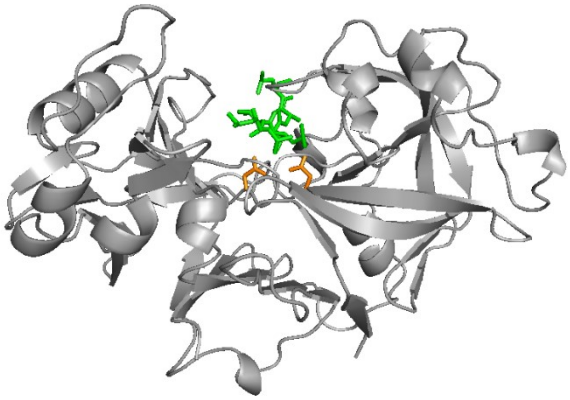
Aspartátové proteasy se podílejí na štěpení proteinů ve fyziologických i patofyziologických procesech a jejich zástupci byli popsáni u virů, bakterií, hub, rostlin i živočichů[9]. U obratlovců se podílejí na katabolickém štěpení proteinů v lysozomech, při trávení potravy ve střevě a jsou důležité pro udržování homeostáze v buňkách[10]. Některé druhy parazitárních členovců nebo prvoků využívají aspartátové proteasy k trávení potravy

získané z hostitele[11,12]. U rostlin jsou aspartátové proteasy nezbytné pro rozmnožování, při obraně proti patogenům, během rostlinné senescence a apoptózy[13,14]. K trávení hmyzu využívají masožravé rostliny aspartátové proteasy[15]. U hub jsou sekretované aspartátové proteasy důležité pro jejich výživu a také při patogenezi[16]. U retrovirů se aspartátové proteasy podílejí na štěpení polyproteinů před složením nové virové částice uvnitř hostitelské buňky[17].

Eukaryotické aspartátové proteasy rodiny A1 pepsinu jsou endopeptidasy, které jsou aktivní převážně v kyselém pH. K nejznámějším zástupcům patří pepsin A, pepsin B, pepsin C a chymosin z trávicího traktu savců. Katepsin D (KatD) a katepsin E jsou především intracelulární proteasy, které se podílejí na proteolýze uvnitř lysozomů[18]. Renin je proteasa sekretovaná z ledvin do krevního řečiště, kde štěpí angiotensinogen na angiotensin I, účastní se tak regulace krevního tlaku. Na rozdíl od většiny aspartátových proteas má renin pH optimum v neutrální oblasti okolo pH 7[19]. β-sekretasa 1 (BACE1) je proteasa vázaná na buněčnou membránu neuronů. BACE1 je schopna štěpit amyloidní prekurzorový protein (APP) na β-amyloidní peptid (β-AP), který se hromadí v amyloidních placích v nervových buňkách při Alzheimerově chorobě[20,21].

1.1.2.2 Struktura aspartátových proteas rodiny pepsinu

Aspartátové proteasy rodiny A1 pepsinu se skládají ze dvou velmi podobných domén označovaných jako C- a N-doména, které pravděpodobně vznikly duplikací jednoho genu[22]. Na obr. 2 na str. 7 jsou znázorněné dvě domény lidského pepsinu. Aktivní centrum se nachází mezi oběma doménami, přičemž jeden katalytický zbytek kyseliny asparagové je v každé doméně. Podlouhlá dutina aktivního centra (o délce cca 30 Å) je schopná vázat až 8 aminokyselin substrátu. Aminokyselinová sekvence v oblasti katalytických aspartátů je konzervovaná, nejčastěji Asp-Thr/Ser-Gly-Thr/Ser. Proteasy rodiny pepsinu sdílejí vysokou podobnost aminokyselinové sekvence i terciární struktury[10].



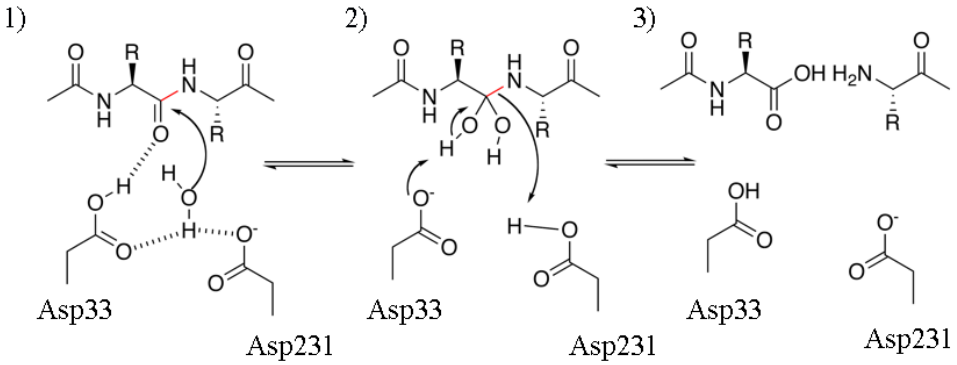
Obr. 2 Prostorová struktura pepsinu v komplexu s inhibitorem pepstatinem A (PDB kód: 1PSO). Molekula pepsinu (šedivě) je znázorněna ve stužkovém modelu. Dva katalytické aspartáty (oranžově) a pepstatin A (zeleně) v aktivním centru jsou zobrazeny v tyčkovém modelu. Obrázek byl připraven pomocí programu Pymol.

Některé aspartátové proteasy mají přidané domény se specifickou funkcí. Například BACE1 má hydrofobní transmembránovou doménu nebo phytepsin, aspartátová proteasa z vakuol ječmene, má na C-konci specifickou doménou PSI („plant specific insert“) o velikosti 10 kDa, která je příbuzná se saposiny a umožňuje interakci s membránami[23,24].

Posttranslační modifikace se vyskytují u všech aspartátových proteas rodiny pepsinu. Tyto proteasy mají tři konzervované disulfidické můstky, které stabilizují terciární strukturu[10]. Další možné posttranslační modifikace jsou glykosylace (např. lidský katepsin E) nebo fosforylace (např. hovězí KatD).

1.1.2.3 Katalytický mechanismus aspartátových proteas

Aspartátové proteasy mají podélné aktivní centrum, na jehož dně jsou dva katalytické zbytky kyseliny asparagové (Asp33 a Asp231). Jeden katalytický aspartát se chová jako kyselina, druhý jako báze, což je dáno rozdílnou hodnotou pKa. Molekula vody, která interaguje s katalytickými aspartáty a je jimi deprotovaná, zahajuje proteolytické štěpení substrátu. Její kyslík jako nukleofil atakuje karbonylový uhlík substrátu (krok 1 na obr. 3 na str. 8) za vzniku nestabilního tetrahedrálního meziproduktu (krok 2 na obr. 3 na str. 8). Přesunem hydroxylu z vody na substrát vzniká nová karboxylová skupina na C-konci produktu, vodík z molekuly vody je součástí nové aminoskupiny na N-konci druhého produktu (krok 3 na obr. 3 na str. 8). Následně dochází k uvolnění obou produktů z dutiny aktivního centra a nakonec se katalytické aspartáty regenerují navázáním nové molekuly vody[25].



Obr. 3 Schematické znázornění katalytického mechanismu aspartátových proteas. Štěpená peptidová vazba je znázorněna červeně. 1) Molekula vody navázaná na katalytické aspartáty (Asp33 a Asp231) zahajuje nukleofilní atak na substrát. 2) Vzniká nestabilní tetrahedrální intermediát. 3) Následuje rozštěpení peptidové vazby za vzniku karboxylové skupiny a amino skupiny produktů. Převzato a upraveno podle[26].

1.1.2.4 Substrátová specifita aspartátových proteas rodiny pepsinu

Všechny aspartátové proteasy rodiny pepsinu jsou endopeptidasy, ale jednotliví zástupci mají různé fyziologické substráty. Například renin má vyhraněnou substrátovou specifitu pouze na jeden známý fyziologický substrát, v krevní plazmě štěpí angiotenzinogen na angiotenzin I[27]. Naopak pepsin, hlavní trávicí proteasa v žaludku, má velmi širokou substrátovou specifitu[28].

Substrátová specifita aspartátových proteas byla studována pomocí knihoven substrátů obsahující různé aminokyseliny v jednotlivých podmístech. Většina proteas rodiny pepsinu preferuje štěpení mezi velkými hydrofobními aminokyselinami[29]. Rozdílnou substrátovou specifitu proteas lze vysvětlit odlišnou stavbou jejich aktivního centra a okolí. Například tyrosinová a polyprolinová smyčka KatD, které se nacházejí nad aktivním centrem, mají vliv na substrátovou specificitu enzymu. Tyrosinová smyčka (Tyr137 a Tyr139) určuje substrátovou specificitu S_1 podmísta. Polyprolinová smyčka se pravděpodobně podílí na vstupu substrátu do aktivního centra[30].

Na základě porovnání preferencí pro aminokyseliny v jednotlivých podmístech aktivního centra byl navržen oktapeptid se sekvencí Lys-Pro-Ile-Glu-Phe↓Nph-Arg-Leu, který je velmi dobrým substrátem pro kinetická měření různých aspartátových proteas[31]. Štěpení vazby mezi Phe↓Nph (Nph – p-nitrofenylalanin) lze snadno detekovat jako pokles absorbance při 305nm[32].

1.1.2.5 Regulace aktivity aspartátových proteas rodiny pepsinu

Regulace aktivity aspartátových proteas je komplexní proces, který probíhá na několika úrovních zahrnujících genovou expresi, translaci proteinu, aktivaci proenzymu a modulaci aktivity zralého enzymu. V této kapitole budou podrobně popsány poslední dva mechanismy.

Aktivace aspartátových proteas rodiny pepsinu

Aspartátové proteasy jsou syntetizované ve formě neaktivního proenzymu (také zvaného zymogen). Aktivace proenzymu probíhá pouze za určitých podmínek a při správné kompartmentaci dané proteasy a může být intracelulární (např. prokathepsin D (proKatD) se aktivuje v endo/lysozomech) nebo extracelulární (např. pepsinogen se aktivuje po sekreci do žaludku). Proenzym má na svém N-konci propeptid (u aspartátových proteas rodiny A1 pepsinu je až padesát aminokyselin dlouhý), bývá také označován jako propart, prosegment nebo aktivační peptid. V této práci budou aminokyseliny propeptidu číslované nezávisle na zralém enzymu, za číslem aminokyseliny bude přípona „p“ (například sekvence lidského proKatD začíná Leu1p). Propeptid má několik funkcí: napomáhá správnému sbalení proenzymu, zvyšuje stabilitu proteinu v buněčném prostředí a reguluje aktivitu proteasy tím, že brání ve vazbě substrátu do aktivního centra. Změna pH je důležitý faktor, který se podílí na řízení procesu aktivace u proteasy[33–35].

Propeptid může být odštěpen celý v jednom kroku, nebo postupně ve více krocích za vzniku aktivačních intermediátů. Aktivace pepsinogenu A probíhá v jednom kroku, zatímco při aktivaci progastricinu vzniká intermediát[36,37]. K odštěpení propeptidu za vzniku katalyticky aktivní proteasy dochází jak transaktivací, tak i autoaktivací. Při transaktivaci jiná, katalyticky aktivní proteasa jiného typu odštěpí propeptid proenzymu. Příkladem trans-aktivace je aktivace proenzymu kvasinkové proteinasy B kvasinkovou proteinasou A[38]. Při autoaktivaci si propeptid odštěpí sama proteasa, tento mechanismus aktivace se dále dělí na monomolekulární a bimolekulární. Monomolekulární autoaktivace probíhá např. u králičího pepsinogenu A-3[36]. Při bimolekulárním mechanismu autoaktivace jiná, již aktivovaná molekula odštěpí propeptid proenzymu tak jako např. u opičího pepsinogenu A[39]. U některých proteas (např. hovězího prochymosinu) může docházel k autoaktivaci oběma způsoby (monomolekulárně i bimolekulárně) v závislosti například na pH prostředí nebo koncentraci enzymu[40].

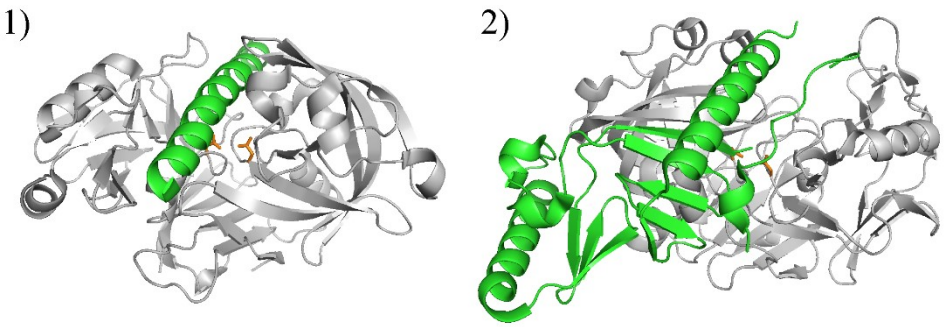
Modulace aktivity aspartátových proteas rodiny pepsinu

I pro katalyticky aktivní enzymy zůstává pH jedním z mechanismů regulace jejich aktivity. Většina aspartátových proteas má kyselé pH optimum. Například pepsin štěpí substrát v oblasti pH 1 – 6, s nejvyšší aktivitou v pH 3,5[41]. Při dalším růstu pH dochází nejdřív ke změně sekundárních struktur a ionizaci některých kyselých aminokyselin, později k denaturaci celé molekuly[42]. U proteas s kyselým pH optimem dochází často při zvýšení pH nad 7 k ireverzibilní denaturaci[43].

Pro aspartátové proteasy byly popsány peptidové i proteinové inhibitory. Nejznámější peptidový inhibitor aspartátových proteas je pepstatin A, hexapeptid z bakterií *Actinomyces*, který je pikomolárním kompetitivním inhibitorem pepsinu[44]. V jeho sekvenci Iva_1-Val_2-Val_3-Sta_4-Ala_5-Sta_6 je kyselina isovalerová (Iva) a nestandartní aminokyselina statin ((3S,4S)-4-amino-3-hydroxy-6-methylheptanová kyselina) (Sta). Při vazbě pepstatinu do aktivního centra pepsinu vytváří hydroxylová skupina Sta_4 inhibitoru vodíkové můstky s katalytickými aspartáty proteasy. Vytěsňuje tak z aktivního centra molekulu vody, která je nezbytná pro zahájení proteolýzy[45]. Prostorová struktura komplexu pepsin-pepstatin A je znázorněna na obr. 2 na str. 7.

Jsou známé i další přirozené peptidové inhibitory aspartátových proteas. Hydrofilní peptid z bakterií rodu *Bacillus* ATBI („alkalo-thermophilic *Bacillus* inhibitor“) je kompetitivním inhibitorem pepsinu i proteasy F-prot z houby *Aspergillus saitoi*[46]. Další známé peptidové inhibitory aspartátových proteas představují jejich propeptidy[47]. U lidského KatD byly popsány dvě oblasti propeptidu, které inhibují aktivitu enzymu[48]. N-koncová část propeptidu se váže pomocí elektrostatických interakcí a C-koncová část propeptidu s tzv. „lysin-tyrosinovou kotvou“ přímo interaguje s katalytickými aspartáty.

Inhibitory aspartátových proteas proteinového charakteru jsou vzácné a mechanismus interakce s enzymem byl popsán pouze u dvou z nich. Prvním takovým inhibitorem je IA3 proteinový inhibitor o hmotnosti 8 kDa z kvasinky *Saccharomyces cerevisie*, který specificky inhibuje aspartátovou proteinasu A[49]. Zatímco v roztoku nemá inhibitor pravidelnou sekundární strukturu, v komplexu s pepsinem tvoří jeden dlouhý α-helix. Druhým je inhibitor PI3 (z angl. „pepsin inhibitor-3“) o hmotnosti 16 kDa z hlístice *Ascaris suum*. Jedná se o specifický proteinový inhibitor pepsinu, který váže svůj N-konec do S1´-S3´ vazebných podmíst enzymu, a tak blokuje vazbu substrátu[50]. Prostorové struktury komplexů proteinasa A-IA3 a pepsin-PI3 byly vyřešeny a jsou znázorněny na obr. 4 na str. 11.



Obr. 4 Prostorové struktury komplexů proteinových inhibitorů s aspartátovými proteasami. Aspartátová proteasa je znázorněna ve stužkovém modelu šedivě, katalytické aspartáty jsou v tyčkovém modelu oranžově. Inhibitory jsou ve stužkovém modelu zeleně 1) Proteinasa A v komplexu s proteinovým inhibitorem IA3 (PDB kód: 1DPJ), 2) Pepsin v komplexu s proteinovým inhibitorem PI3 (PDB kód: 1F34). Obrázek byl připraven pomocí programu Pymol.

U ostatních proteinových inhibitorů aspartátových proteas zatím nebyl popsán přesný mechanismus inhibice. Z lilku brambor byl izolován specifický inhibitor pro KatD nazvaný „Potato Cathepsin D Inhibitor“ (PDI)[51]. Jedná se o 23 kDa glykoprotein, který patří do rodiny Kunitzových inhibitorů. PDI je bifunkční inhibitor, který je schopen účinně blokovat aktivitu aspartátových i serinových proteas[52,53]. α2-makroglobulin, protein z plazmy obratlovců, inhibuje řadu aspartátových proteas jako jsou chymosin a katepsiny D a E, ale také serinové, cysteinové i metaloproteasy[54]. Inhibitor „Squash Aspartic Protease Inhibitor“ (SQAPI), protein s 96 aminokyselinami z dýně *Cucurbita maxima*, je sekvenčně podobný rostlinným inhibitorům cysteinových proteas tzv. cystatinům, ze kterých pravděpodobně vznikl[55]. Jedná se o nanomolární inhibitor pepsinu a sekretované aspartátové proteasy z houby *Glomerella cingulata*. Tento inhibitor pravděpodobně neinteraguje s katalytickými aspartáty proteasy a tím je chráněn před degradací[55,56]. Protein equistatin o velikosti 22 kDa z mořské sasanky *Actinia equina*, je tvořen třemi doménami tyreoglobulinového typu-1. Doména 2 účinně inhibuje KatD, jiné aspartátové proteasy ovšem neinhibuje[57,58]. Doména 1 equistatinu inhibuje cysteinové proteasy, funkce třetí domény zatím není známá[57].

Syntetické inhibitory aspartátových proteas byly připraveny na základě znalosti sekvence přirozených inhibitorů. Například na základě struktury inhibitoru pepstatinu byly odvozeny další peptidomimetické inhibitory aspartátových proteas, které obsahují Sta[59]. Další strategie pro návrh inhibitorů aspartátových proteas je založena na základě známých substrátů těchto enzymů. Tyto inhibitory napodobují přechodný stav během proteolýzy, mají ale pozměněnou strukturu tak, aby vazba mezi P_1-P_1‘ zbytky nebyla hydrolyzovatelná. Navrženy byly například inhibitory s redukovanou amidovou vazbou,

s hydroxyethylenovým isosterem nebo se záměnou L-aminokyselin za D-aminokyseliny[60–62].

Další skupinou inhibitorů aspartátových proteas jsou látky, které kovalentně modifikují katalytické aspartáty. Po jejich navázání už katalytické aspartáty nemohou interagovat s molekulou vody a následně zahájit proteolýzu. Jako příklad lze uvést inhibitor EPNP (1,2-epoxy-3-(p-nitrophenoxy)propan), který esterifikuje katalytické aspartáty nebo peptidomimetické inhibitory s epoxyalkoholovou skupinou, které alkylují katalytické aspartáty[63,64]. Všechny inhibitory modifikující katalytické aspartáty jsou ale nespecifické, protože mohou modifikovat i jiné aminokyseliny v molekule enzymu.

1.1.3 Proteasy typu katepsin D

Název katepsiny pochází z řeckého „kathepsein“, česky „trávit“. Termín katepsin D použil poprvé Press v roce 1960 pro označení proteasy izolované z hovězí sleziny s kyselým pH optimem[65]. V současné době je známo dvacet různých katepsinů, v lidském genomu jich je popsáno patnáct. Nejvíce zastoupené jsou cysteinové katepsiny, například katepsiny B, C, H, L. Katepsiny A a G jsou serinové proteasy a katepsiny D a E jsou aspartátové proteasy[66].

1.1.3.1 Funkce proteas typu katepsin D

Experimenty na myších s vypnutým genem pro KatD prokázaly důležitost KatD pro správný vývoj. Během embryogeneze KatD pravděpodobně není nezbytný, na myších nebyly pozorovány žádné změny ani během prvních dvou týdnů po narození[67]. Je ale možné, že KatD přecházel z matky do embrya přes placentu během prenatálního vývoje a po porodu v mléku. U myší během třetího týdne po narození docházelo k rozsáhlé nekróze v tenkém střevě a k souvisejícímu úbytku hmotnosti. Apoptóza byla pozorovaná u buněk oční sítnice, kde vedla k oslepnutí, a také u buněk v brzlíku a ve slezině. Změny byly pozorovány i v centrálním nervovém systému, kde docházelo k hromadění ceroidního lipofuscinu v neuronech podobně jako u neuronální ceroidní lipofuscinózy[68]. Proteolýza v lysozomech byla zachovaná v obvyklém rozsahu, je ale možné, že ostatní katepsiny kompenzovaly absenci KatD[69]. Nicméně všechny myši uhynuly čtvrtý týden po narození[67].

Dalšími modelovými organismy, u kterých byl odstraněn gen pro katepsin D, byly ryby a mouchy octomilky (*Danio rerio* a *Drosophila melanogaster*). U ryb docházelo k abnormálnímu vývoji očí, hyperpigmentaci kůže, nevyvíjel se u nich plynový měchýř a celkově byly menší a dříve uhynuly[70]. Octomilky se vyvíjely normálně a byly

životaschopné a fertilní, nicméně i u nich docházelo k neurodegeneraci podobné neuronální ceroidní lipofuscinóze[71]. Výsledky studií na modelových organismech prokazují komplexní působení KatD na více úrovních.

U bezobratlých jsou proteasy typu KatD důležité při trávení potravy. Podrobněji je prozkoumaná role KatD u krev sajících parazitů, kde KatD ve střevech zahajuje štěpení hemoglobinu, důležitého zdroje energie a aminokyselin. KatD z parazitické motolice krevničky *Schistosoma japonicum* (SjKatD1), je lokalizován v trávicích vakuolách a pravděpodobně je důležitý během štěpení hemoglobinu[72]. Krevnička *Schistosoma mansoni* má tři různé proteasy typu KatD – SmKatD1, SmKatD2 a SmKatD3. SmKatD1 zahajuje štěpení v hemoglobinolytické kaskádě ve střevech[73]. SmKatD2 byl nalezen v dospělých parazitech, SmKatD3 v larvách schistosomulách[12]. Inhibice SmKatD1 pomocí RNA interference vedla ke ztrátě schopnosti krevničky trávit hemoglobin a v důsledku k supresi růstu krevničky[74]. Ve střevech ploštice *Triatoma infestans* byly nalezeny dvě proteasy typu KatD - TiKatD1 a TiKatD2. Exprese TiKatD1 byla závislá na sání krve z hostitele[75].

U řady bezobratlých, kteří nepatří mezi parazity, je KatD také přítomný ve střevech jako trávicí proteasa. U mouchy domácí (*Musca domestica*) se proteasy typu KatD podílejí na trávení ve předním a středním střevě[76]. Ve střevech zrnokaze skvrnitého (*Callosobruchus maculatus*) a mandelinky bramborové (*Leptinotarsa decemlineata*) proteasy typu KatD spolu s cysteinovými proteasami štěpí rostlinné proteiny[77,78]. V přítomnosti inhibitorů cysteinových proteas dochází k zvýšené expresi proteas typu KatD, které tyto inhibitory štěpí a zvyšují tak aktivitu cysteinových proteas ve střevě.

U hmyzu se KatD podílí na štěpení proteinu vitelogeninu nebo později vitellinu, zásobního proteinu vajec, a je nezbytný během procesu metamorfózy. U komárů *Aedes aegypti* se lysozomální proteasa typu KatD označovaná mLAP podílí na syntéze a zpracování vitellogeninu[79]. Podobně u ploštic *Rhodnius polixus* a *Dipetalogaster maxima* proteasa typu KatD degraduje vitelin během embryogeneze[80,81]. Proteasa typu katepsin D (BmCatD) z bource morušového (*Bombyx mori*) se účastní programované buněčné smrti střev a tukového tělesa během proměny larvy na dospělce[82].

1.2 Lidský katepsin D

1.2.1 Fyziologie lidského katepsinu D

1.2.1.1 Lokalizace a fyziologické funkce lidského katepsinu D

Lidský KatD je lokalizován především v lysozomech, pozdních endosomech a fagosomech všech tkání a orgánů[83]. Během apoptózy dochází k degradaci lysozomů a vylití jejich obsahu do cytosolu (viz níže). Byl také popsán přesun KatD z lysozomů do buněčného jádra, kde se váže na chromatin[84]. Jeho funkce v buněčném jádře ale není zatím známá.

Exocytózou je KatD sekretován z buněk. U lidí byl KatD nalezen v séru, v potu a v moči[85–87]. V epitelových buňkách mléčné žlázy dochází k selektivní exocytóze proKatD[88]. Dále byl KatD lokalizován v extracelulární matrix buněk, v synoviální tekutině kloubů a v cévách v místech, kde se hromadí makrofágy[89–91].

V lysozomech lidský KatD degraduje nespecificky funkční a strukturní proteiny, ale také inaktivuje peptidové hormony a inhibitory. Dále se podílí na specifické aktivaci růstových faktorů nebo jejich receptorů. Mezi jeho fyziologické substráty patří apolipoprotein B, aktin, myosin, fibrin, inzulin, glukagon, kininogen, cystatin C a mnoho dalších proteinů a peptidů[92]. Produkty štěpení KatD jsou také peptidy, které jsou prezentovány na MHC II receptorech na povrchu imunitních buněk[93].

Programovaná buněčná smrt (apoptóza) zahrnuje sled biochemických procesů, které vedou k řízenému šetrnému odstranění buňky z organismu a lysozomální proteasy jsou důležitou součástí tohoto procesu[94]. Zvýšená propustnost lysozomální membrány je jeden z prvních kroků apoptózy a předchází například rozpadu mitochondrií, uvolnění cytochromu c nebo aktivaci caspas. KatD se účastní různých apoptotických drah a může indukovat apoptózu přímou nebo nepřímou dráhou. Je mediátorem apoptózy při oxidativním stresu nebo po přidání cytotoxických látek, např. staurosporinu[95,96].

Nedostatečně prozkoumaná je schopnost KatD štěpit proteiny v cytosolu, kde je fyziologické pH okolo pH 7,2. Během apoptózy dochází k okyselení cytosolu, což je příznivé pro katalytickou aktivitu KatD, který má kyselé pH optimum[97]. V cytosolu fibroblastů štěpí KatD protein Bid (mediátor apoptózy) na třech místech, což vede k zahájení apoptózy[98]. Dále je katalyticky aktivní KatD nezbytný pro přesun Bax proteinu na mitochondriální membránu, kde dochází po oligomerizaci Bax proteinů k zahájení

apoptózy[99,100]. V cytosolu endoteliálních buněk KatD štěpí thioredoxin-1, důležitý anti-apoptotický protein[101]. Podle některých studií během apoptózy není nezbytná katalytická aktivita KatD. Přidáním KatD (katalyticky aktivního i neaktivního) do fibroblastů nebo nádorových HeLa buněk vyvolává v obou případech apoptózu[102]. Rozvoj apoptózy v nádorových buňkách jako odpověď na přítomnost protinádorových léčiv je mnohem rychlejší v přítomnosti KatD, nezávisle na jeho katalytické aktivitě [103].

Na druhou stranu podle jiných studií především v nádorových buňkách může zvýšená exprese KatD podporovat růst nádoru a inhibovat apoptózu, tj. působit anti-apoptoticky. U nádorových i nenádorových buněk exprese KatD stimuluje proliferaci buněk a angiogenezi[104]. Transfekce KatD do fibroblastů podporuje jejich růst nezávisle na aktivitě enzymu[105]. Naopak inhibice KatD v neuroblastomech specifickým inhibitorem pepstatinem vedla v některých studiích k indukci apoptózy činností caspas[100]. Z výše uvedených studií je patrné, že fyziologické a patofyziologické působení KatD je komplexní a doposud nebylo plně objasněno.

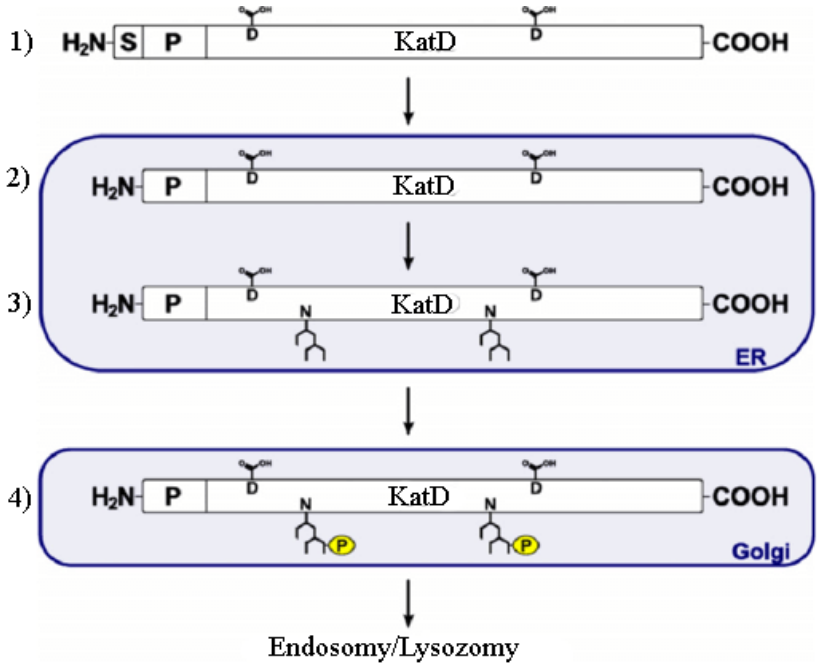
1.2.1.2 Syntéza a aktivace lidského katepsinu D

Gen pro KatD je lokalizován na chromozomu 11, skládá se z 9 exonů[106]. Promotor obsahuje sekvence typické pro konstitutivní expresi (bohaté na CG a sekvence pro vazbu transkripčního faktoru Sp1) i charakteristické pro inducibilní expresi (TATAA sekvence) [107]. Lze tedy předpokládat, že dochází k oběma typům exprese. Exprese KatD je regulovaná růstovými faktory, cytokiny, steroidními hormony nebo vitaminy, například estrogen a progesteron zvyšují jeho expresi[108–110].

Lidský KatD je syntetizován ve formě pre-pro-proteinu na ribozomech hrubého endoplasmatického retikula, obsahuje 20 aminokyselin dlouhý signální peptid, následuje 44 aminokyselin dlouhý propeptid a dále vlastní enzym o délce 348 aminokyselin. Signální peptid je odštěpen signální peptidasou během přesunu přes membránu do lumen hrubého endoplasmatického retikula[111].

Uvnitř hrubého endoplasmatického retikula probíhá glykosylace proenzymu, na dva zbytky asparaginu (Asn70 a Asn199) jsou navázány oligosacharidy vysokomannosového typu[112]. Tato glykosylace není důležitá pro enzymovou aktivitu ani správné sbalení proteinu, ale pro jeho další směrování do lysozomů[113]. V *cis*-Golgiho aparátu jsou mannosové zbytky fosforylovány. Tento krok je nezbytný aby se proKatD v *trans*-Golgi aparátu navázal na mannosa-6-fosfátové receptory (M6F-R) a dostal se do endo-

lysozomálních kompartmentů, kde se proKatD působením kyselého pH uvolňuje z M6F-R[114]. Schéma syntézy a kompartmentace proKatD jsou znázorněné na obr. 5 na str. 16.



Obr. 5 Schematické znázornění syntézy a transportu proenzymu lidského prokatepsinu D (proKatD). 1) Syntéza pre-pro-enzymu KatD (S – signální sekvence, P – propeptid) probíhá na ribozomech hrubého endoplasmatického retikula (ER). 2) Signální peptid je odštěpen během přesunu do lumen endoplasmatického retikula signální peptidasou. 3) V ER dochází ke glykosylaci vysokomannosového typu na dvou asparginylových zbytcích (N). 4) Následná fosforylace mannosových zbytků (P ve žlutém kolečku) probíhá v Golgiho aparátu (Golgi), kde se takto upravený proKatD váže na M6F-R a tento komplex je následně transportován do endo/lysozomů. Schematicky jsou znázorněny katalytické aspartáty Asp33 a Asp231 (D). Převzato a upraveno podle[115].

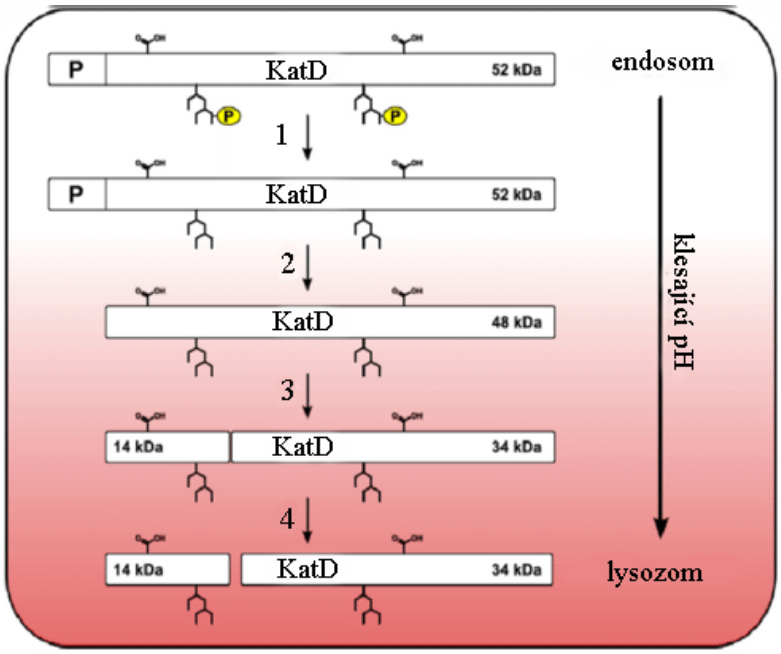
K transportu do lysozomů může docházet i nezávisle na M6F-R. Saposiny A, B, C a D jsou malé proteiny, které aktivují v lysozomech hydrolázy sfingolipidů. Prosaposiny (vysokomolekulární prekurzor saposinů) se mohou vázat na proKatD za vzniku stabilního komplexu oligomerů prosaposinů s proKatD[116]. Vzniklý komplex může být transportován do lysozomů nebo být sekretován ven z buňky[116,117]. V lysozomech prosaposin stimuluje aktivaci proKatD na katalyticky aktivní pseudokatepsin D (pseudoKatD), který následně štěpí prosaposiny na saposiny[116,118]. Uvedený mechanismus nezávislý na M6F-R je specifický a funkční pro některé buněčné linie bez M6F-R[119].

V kyselém pH endo/lysozomů se nejprve proKatD uvolňuje z M6F-R, následně dochází k odštěpení fosfátů ze sacharidových zbytků na povrchu enzymu (krok 1 na obr. 6 na str. 18) a poté k samotné aktivaci enzymu. V prvním kroku dochází vlivem pH k rozvolnění vazby mezi propeptidem a enzymem, což umožní odštěpení N-konce propeptidu autoaktivací. Během tohoto monomolekulárního procesu je odstraněno prvních 26 aminokyselin z propeptidu (Leu1p-Leu26p) za vzniku katalyticky aktivní formy

pseudoKatD[120–122]. Tato forma KatD je schopná aktivovat další molekuly proKatD a urychlit tak vznik aktivního enzymu bimolekulárním mechanismem. Autoaktivace proKatD na pseudoKatD byla potvrzena pouze *in vitro*, pseudoKatD nebyl izolován *in vivo* a tento krok není ani nezbytný pro vznik zralého enzymu[123]. V dalším kroku aktivace dochází k odštěpení buď zbytku propeptidu (Leu27p-Glu44p) nebo celého propeptidu (Leu1p-Glu44p) jinou proteasou (krok 2 na obr. 6 na str. 18). Tento proces je bimolekulární, vzniká zralý jednořetězcový KatD (sekvence začíná Gly1)[122,124]. Aktivační proteasou je pravděpodobně některá cysteinová proteasa, katepsin B nebo katepsin L[125,126].

Dále je v lysozomech vyštěpen z jednořetězcového KatD hexapeptid Ala99-Ala104, čímž vzniká dvouřetězcová forma KatD, složená z lehkého řetězce na N-konci (14kDa) a těžkého řetězce na C-konci (34kDa) (krok 3 na obr. 6 na str.18)[127]. Lehký a těžký řetězec nejsou spojené kovalentně. Lidský KatD se téměř výhradně vyskytuje ve dvouřetězcové formě a není známý důvod změny jednořetězcové formy KatD na dvouřetězcovou formu[128].

Popsáno bylo zkracování koncových sekvencích zralého KatD. Karboxypeptidasy nebo aminopeptidasy mohou zkrátit jeden nebo oba řetězce až o sedm aminokyselin postupným odbouráváním (krok 4 na obr. 6 str. 18)[123,129–131]. Předpokládá se, že postupné odbourávání aminokyselin z N- a C-konce KatD souvisí se stárnutím molekuly[130]. Další popsaná modifikace KatD je nitrace tyrosinového zbytku Tyr168, která byla popsaná při remodelaci prsní tkáně a vedla ke zvýšení enzymové aktivity[132].



Obr. 6 Schéma aktivace a procesování proenzymu lidského katepsinu D (proKatD) za vzniku aktivního enzymu. 1) V endosomech jsou sacharidové zbytky proKatD (52 kDa) defosforylovány (P ve žlutém kolečku). 2) Odštěpení propeptidu (P) za vzniku zralého jednořetězcového KatD (48 kDa). 3) Dále je tato forma rozštěpená na dvouřetězcovou formu (N-koncová doména o velikosti 14k Da a C-koncová doména o velikosti 34 kDa). 4) Postupné odbourávání aminokyselin z N- a C-konců obou řetězců KatD působením aminopeptidas a karboxypeptidas. Převzato a upraveno podle[115].

1.2.2 Patofyziologie lidského katepsinu D

Lidský KatD je spojován s celou řadou patologií, mezi nejznámější patří nádorová onemocnění, Alzheimerova choroba a ateroskleróza, mezi ty méně známé se řadí například osteoartritida, akutní zánět slinivky břišní nebo kožní onemocnění.

1.2.2.1 Katepsin D v nádorových onemocněních

Zvýšené množství neaktivního proKatD v nádorových buňkách různých tkání bylo poprvé pozorováno ve druhé polovině 80. letech[133]. Od té doby řada studií popsala zvýšený výskyt proKatD a/nebo KatD v celé řadě nádorových onemocnění. V současné době je zvýšená hladina proKatD považovaná za nezávislý prognostický marker nádorového onemocnění prsu[134,135]. Nicméně kvůli nejednotné metodice kvantifikace proKatD není takto využíván[134].

U nádorů prsu, prostaty, ovarií nebo plic byla popsána nadprodukce proKatD a jeho následná sekrece do extracelulárního prostoru[136–139]. Nádorové buňky, které sekretují laktát a ovlivňují funkci H+/ATPásové pumpy na povrchu buněčné membrány, tvoří ve svém okolí mírně kyselé pH, které je vhodné pro aktivaci proKatD na katalyticky aktivní KatD,

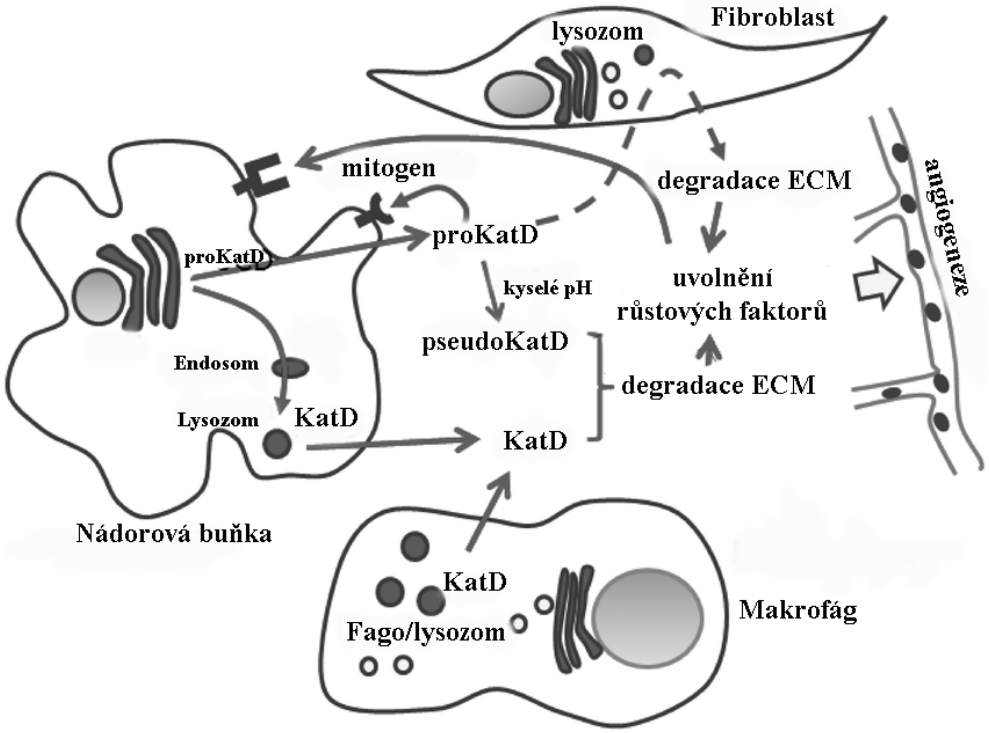
který by mohl dále štěpit extracelulární matrix (ECM) a rozpustné extracelulární proteiny, například chemokiny, které se podílejí na protinádorové imunitě, inhibitory jiných proteas, které štěpením ECM mohou usnadňovat růst nádoru, nebo přímo růstové faktory [140–143].

Jiné studie prokazují autokrinní aktivitu KatD nezávislou na jeho katalytické aktivitě[144]. Po přidání pepstatinu, inhibitoru aspartátových proteas, nebyl pozorován žádný vliv na změnu růstu nádorových buněk[145]. Dále byla popsaná vazba proKatD na buněčnou membránu nádorových buněk prsu, kterou bylo možné blokovat specifickými protilátkami proti propeptidu proKatD[146,147]. Vazba syntetického propeptidu na povrch nádorových buněk (například u ovarií, prostaty, prsu nebo plic), měla stejný mitogenní účinek jako vazba celého proenzymu[148]. Jiné studie uvádějí interakci těžkého řetězce KatD s doposud nepopsaným povrchovým receptorem, která byla nezávislá na katalytické aktivitě enzymu[149].

Doposud byly popsány dva receptory, na které by se proKatD mohl vázat – M6F-R a membránový receptor „low-density lipoprotein receptor-related protein-1“ (LRP1), oba se nachází na buněčné membráně. Vazba proKatD na β-řetězec receptoru LRP1 na povrchu fibroblastů vede k parakrinní stimulaci růstu fibroblastů[150]. Zatím není jednoznačné, zda LRP1 je receptor, na který se váže proKatD v národových buňkách, nebo jestli se váže na jiný, doposud neznámý receptor.

Pro proKatD bylo popsáno několik různých mechanismů, kterými podporuje šíření nádorových buněk a růst metastází. U fibroblastů proKatD podporuje jejich proliferaci, přežití a pohyblivost, což může mít vliv na další rozrůstání nádoru[144]. U nádorového onemocnění prsu je přítomnost KatD v extracelulárním prostoru spojena s rychlejším rozrůstáním nádoru a snížením kontaktní inhibice buněk[151]. Nepřímé působení KatD na stimulaci růstu metastází je štěpení inhibitorů jiných extracelulárních proteas, které poté mohou degradovat ECM[152]. Štěpením některých chemokinů makrofágů (např. MIP-1, CCL3, CCL4 a SCL21) může KatD také snižovat protinádorovou imunitu a ovlivňovat migraci nádorových buněk[153].

Dalším důležitým procesem, kterého se KatD účastní, je angiogeneze, tj. tvorba nových krevních kapilár, které vyživují nádorové buňky. Množství KatD koreluje s hustotou kapilár v nádorech prsu, ať je enzym katalyticky aktivní nebo není[104]. V *in vitro* testech bylo prokázáno, že KatD je schopen štěpit angiogenní faktory (např. „fibroblast growth factor“ - FGF) i anti-angiogenní faktory (např. angiostatin, prolaktin a endostatin) [142,154–157]. Není ale jasné, které z těchto faktorů je KatD schopen štěpit také *in vivo*[158]. Komplexní působení KatD v nádorové tkáni je souhrnně znázorněno na obr. 7 na str. 20.



Obr. 7 Souhrn funkcí katepsinu D (KatD) v nádorových tkáních. Prokatepsin D (proKatD) je nadměrně syntetizován a následně sekretován z nádorových buněk. V kyselém extracelulárním prostředí dochází k jeho aktivaci. Zralý KatD může být sekretován nejen z nádorových buněk, ale i z makrofágů nebo fibroblastů. Katalyticky aktivní enzym v extracelulárním prostoru může štěpit nejen extracelulární matrix (ECM), ale i uvolnit růstové faktory. Ty, spolu s proKatD, mohou stimulovat další růst nádorových buněk. Růstové faktory stimulují dělení buněk a angiogenezi. Převzato a upraveno podle[159].

Lidský KatD hraje důležitou roli při vzniku a rozvoji nádorových onemocnění, proto je perspektivním cílem pro léčbu některých nádorů[160]. V *in vitro* testech inhibitory KatD zpomalovaly růst vybraných typů nádorů prsu[161]. V současné době jsou navrhovány nové specifické inhibitory, které bude možné využít pro další vývoj protinádorových léčiv[162].

1.2.2.2 Neurodegenerativní onemocnění spojovaná s katepsinem D

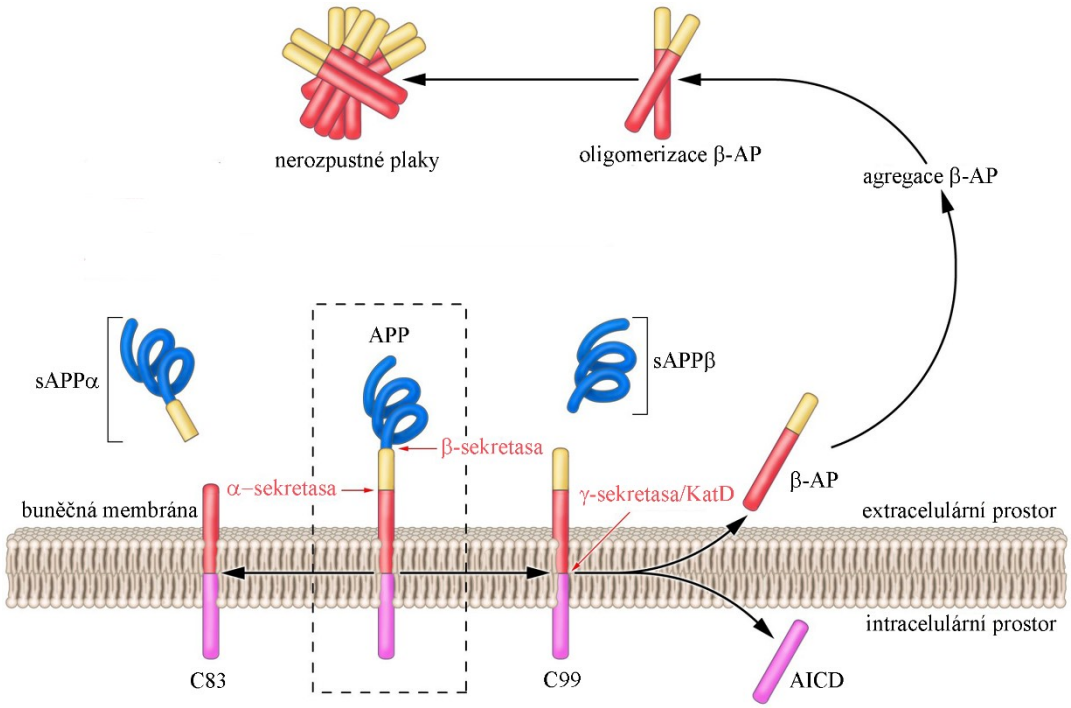
Podobně jako v ostatních tkáních i v neuronech se KatD podílí na degradaci proteinů. Mezi jeho fyziologické substráty patří proteiny jako APP, α-synuclein nebo huntingtin, jejichž chybná degradace je příčinou vzniku různých degenerativních onemocnění (Alzheimerova, Parkinsonova nebo Huntingtonova choroba)[163]. Jednou z možných příčin těchto onemocnění je chybná degradace proteinů KatD. Mutace v sekvenci KatD byly popsány u pacientů s neuronální ceroidní lipofuscinosou, vážným

neurodegenerativním onemocněním[163]. Tato kapitola bude dále zaměřena pouze na Alzheimerovu chorobu, protože s tou je KatD spojován nejčastěji.

Při Alzheimerově chorobě, která je hlavní příčinou stařecké demence, dochází ke ztrátě kognitivních a intelektuálních schopností, ke změnám chování a úbytku fyzických schopností. Během tohoto onemocnění dochází k extracelulárnímu ukládání β-AP a fragmentů z tau proteinů, následnému vzniku plaků na povrchu neuronů a k postupnému odumírání neuronů[164,165].

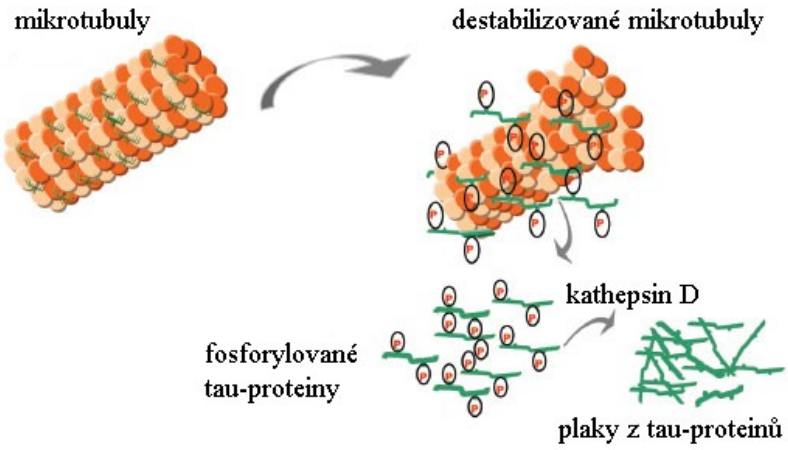
V neuronech se vyskytuje APP o velikosti 695 aminokyselin, který se skládá ze tří domén – extracelulární N-koncové domény, transmembránové domény a intracelulární C-koncové domény. APP je v neuronech fragmentován třemi proteasami (α-, β-, γ-sekretasou). Ve zdravých buňkách je nejprve štěpen α-sekretasou v extracelulární doméně 12 aminokyselin před transmembránovou doménou za vzniku velkého rozpustného fragmentu (sAPPα) a transmembránové domény, která je dále štěpená γ-sekretasou na tzv. p3 peptid. Tímto způsobem nevznikají nerozpustné plaky[166].

Při Alzheimerově chorobě je APP štěpen BACE1 v extracelulární doméně o 16 aminokyselin blíže k N-konci APP než štěpí α-sekretasa. Vzniká tak menší rozpustný fragment (sAPPβ), který se uvolňuje do extracelulárního prostoru a transmembránová doména (C99), která je dále štěpená γ-sekretasou na 28 aminokyselin dlouhý β-AP a protein AICD („APP intracellular domain“). Agregací β-AP v extracelulárním prostoru neuronů vznikají patogenní amyloidní plaky[165]. KatD je schopen štěpit APP ve stejné pozici jako β-sekretasa, proto může přispívat ke vzniku β-AP a patogenních plaků[167]. Zvýšená hladina KatD v mozkomíšním moku a v placích v neuronech byla popsaná u pacientů s Alzheimerovou chorobou[168–170]. Rozdílné štěpení APP α-sekretasou, BACE1 a KatD je schematicky znázorněno na obr. 8 na str. 22.



Obr. 8 Štěpení APP v buněčné membráně sekretasami. APP (v čárkovaném obdélníku) je štěpen α-sekretasou za vzniku extracelulárního proteinu sAPPα a transmembránové domény C83. Při štěpení APP β-sekretasou vzniká transmembránová doména C99. Její následné štěpení γ-sekretasou nebo KatD dochází ke vzniku extracelulárních β-AP a AICD proteinu. Agregací β-AP v extracelulárním prostoru vznikají nerozpustné plaky, které vedou k umírání neuronů. Převzato a upraveno podle[166] .

Další možnou příčinou vzniku Alzheimerovy choroby je tvorba intracelulárních plaků z hyperfosforylovaných tau proteinů, což vede ke zhoršení vnitrobuněčného transportu a komunikaci mezi neurony[164]. Tau proteiny se za fyziologických podmínek nachází v cytosolu neuronů navázané na mikrotubuly, kde napomáhají správnému vnitrobuněčnému transportu. Tau protein má řadu fosforylačních míst, které slouží k regulaci jeho aktivity. Defosforylovaný tau protein brání rychlé depolymerizaci mikrotubulů, zatímco fosforylovaný tau protein depolymerizaci naopak urychluje[171]. KatD je schopen štěpit tau proteiny na C-konci za vzniku neurotoxických fragmentů[172]. Při inhibici KatD v *in vitro* testech dochází k snížení množství těchto fragmentů[173]. Schéma vzniku plaků z tau proteinů uvnitř neuronů je na obr. 9 na str. 23.



Obr. 9 Schematické znázornění vzniku plaků z tau proteinů uvnitř neuronů při Alzheimerově chorobě. Tau proteiny jsou součástí mikrotubulů a jejich fosforylace vede k destabilizaci, rozpadu mikrotubulů a následnému uvolnění tau proteinů do cytosolu. Štěpením tau proteinů v cytosolu KatD mohou vznikat nerozpustné plaky, které způsobují Alzheimerovu chorobu. Převzato a upraveno podle[171].

Dříve byla Alzheimerova choroba spojovaná s minoritním polymorfismem v sekvenci propeptidu proKatD Ala38pVal, který vedl ke změně aktivace proKatD nebo sekreci proKatD do extracelulárního prostoru[48]. Výskyt této mutace je spojen s vyšší pravděpodobností vzniku Alzheimerovy choroby[174]. Nicméně výsledky posledních studií na velkých populacích jsou rozporuplné, proto se soudí, že význam tohoto polymorfismu pro rozvoj Alzheimerovy choroby je spíš malý[175].

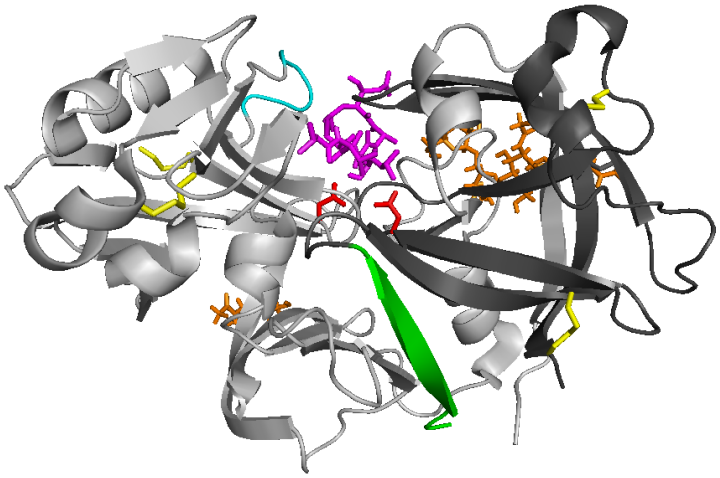
1.2.2.3 Role katepsinu D v ateroskleróze

Ateroskleróza neboli kornatění tepen je onemocnění, při kterém v důsledku ukládání tukových látek do stěn cév dochází ke vzniku plaků a postupnému zúžení až uzávěru cév, což vede k infarktu myokardu nebo mozkové mrtvici[176]. V aterosklerotických placích makrofágy sekretují lysozomální enzymy, mimo jiné proKatD, které se mohou v kyselém prostředí tvořeném makrofágy aktivovat[91,177–180]. Katalyticky aktivní KatD může štěpit LDL („low density lipoproteins“) nebo apolipoprotein B-100, což vede k destabilizaci těchto lipoproteinů, následnému spojování a jejich hromadění ve stěnách cév[91,181]. Makrofágy absorbují změněné LDL a mění se na tzv. pěnové buňky, které se ukládají podél stěn cév a vznikají tak aterosklerotická jádra, základ aterosklerotických plaků[91,177,182].

1.2.3 Biochemická charakterizace lidského katepsinu D

1.2.3.1 Struktura lidského katepsinu D

Lidský KatD má podobnou strukturu jako jiné aspartátové proteasy rodiny pepsinu. Molekula je rozdělená na dvě prostorové domény, které vykazují prostorovou podobnost, a proto se předpokládá, že vznikly duplikací jednoho genu[22]. Mezi doménami se nachází aktivní centrum, přičemž každá doména nese jeden katalytický aspartát (Asp33 a Asp231 číslované podle lidského zralého KatD). V blízkosti aktivního centra se nachází povrchová smyčka se sekvencí bohatou na prolin (Pro312-Pro317), která se podílí na substrátové specificitě enzymu. V molekule jsou celkem čtyři disulfidické můstky, což je o jeden navíc, než mají jiné aspartátové proteasy (disulfidový můstek Cys27-Cys96). Lidský KatD je glykosylovaný oligosacharidovými zbytky vysokomannosového typu v pozicích Asn70 a Asn199[30]. Lidský KatD má vysokou homologii s KatD z jiných savců i s jinými aspartátovými proteasami (např. 87% s prasečím KatD, 49% s pepsinem, 58% s reninem)[183]. Prostorová struktura zralého lidského katepsinu D v komplexu s pepstatinem je znázorněna na obr. 10 na str. 24.



Obr. 10 Prostorová struktura komplexu lidského KatD s inhibitorem pepstatinem (PDB kód: 1LYB). Enzym je zobrazen stužkovým modelem, N-koncový řetězec je tmavě šedý, C-koncový řetězec je světle šedý. Zvýrazněny jsou postranní řetězce katalytických aspartátů Asp33 a Asp231 (červeně), disulfidické můstky (žlutě), oligosacharidové zbytky N-glykosylace (oranžově), sekvence bohatá na prolin Pro312-Pro317 (azurově). N-konec enzymu je označen zeleně. Molekula pepstatinu v aktivním centru je znázorněna v tyčinkovém modelu fialově. Obrázek byl připraven pomocí programu Pymol.

Dvě prostorové struktury lidského KatD byly získány z krystalizačních podmínek v mírně kyselém pH: jednak KatD s prázdným aktivním centrem (PDB kód: 1LYA), jednak KatD v komplexu s nanomolárním inhibitorem pepstatinem. (PDB kód: 1LYB)[184]. U obou

struktur je vidět N-konec enzymu v konformaci jednoho z řetězců β-skládaného listu na povrchu enzymu (na obr. 10 na str. 24 zobrazen zeleně). Dále byla publikována struktura lidského KatD v pH 7.5 (PDB kód: 1LYW), ve kterém je vidět N-konce enzymu přetočený do aktivního centra[185]. Nedávno byly publikovány celkem čtyři struktury KatD s nepeptidovými inhibitory. Tři z nich jsou s lidským KatD (PDB kódy: 4OD9, 4OBZ a 4OC6), jedna je s krysím KatD (PDB kód: 5UX4)[186,187].

Struktura lidského proKatD nebyla zatím vyřešena, publikovaný strukturní model proKatD ukazuje dvě oblasti propeptidu, které interagují s povrchem enzymu. N-koncová doména (Leu1p-Arg14p) se váže k enzymu převážně elektrostaticky. Druhá oblast propeptidu (Ile27p-Glu44p) se váže pomocí tzv. „lysin-tyrosinové kotvy" (Lys34p-Tyr35p) vodíkovými můstky na katalytické aspartáty. Dále bylo pomocí molekulárního modelu prokázáno, že mutace Ala38pVal, spojovaná s Alzheimerovou chorobu, ovlivňuje interakci propeptidu v aktivním centru[48].

1.2.3.2 Proteolytická aktivita lidského katepsinu D

Aktivní centrum lidského KatD je schopné vázat osm aminokyselin (do pozic P_5-P_3'). Pomocí synteticky připravených substrátů bylo zkoumáno, jaké aminokyseliny jsou preferovány v jednotlivých vazebných podmístech enzymu. KatD štěpí substrát mezi dvěma velkými hydrofobními aminokyselinami v pozicích P_1 a P_1'. V pozici P_2' je nejvhodnější bazická aminokyselina (Arg nebo Lys), substrát s Asp v této pozici enzym neštěpí[188,189]. V pozici P_2 KatD preferuje Leu nebo Glu a substrát s Lys v této pozici štěpen není[189,190]. Substrát v aktivním centru enzymu vytváří obvykle strukturu β-skládaného listu umožňující přístup ke katalytickým aspartátům[189].

Lidský KatD má pH optimum typicky v kyselé oblasti mezi pH 3,0 a pH 5,0 podle zvoleného substrátu, například pro hemoglobin bylo určeno pH optimum 3,0, pro peptidové substráty pH 5,0[191,192]. Při vyšším pH dochází k postupnému poklesu aktivity až do pH 6,0, kdy už enzym přestává být aktivní[65]. Některé substráty KatD štěpí i při pH blízkém neutrální oblasti, uváděn je např. cytosolární Bid-protein při pH 6,2 a tau-protein při pH 7,0[193,194].

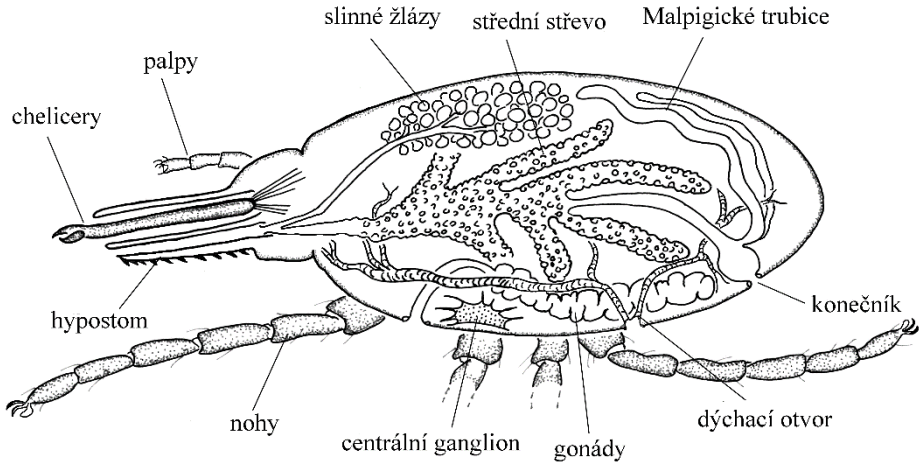
1.3 Katepsin D z parazitických klíšťat

1.3.1 Taxonomie, morfologie a životní cyklus klíštěte obecného

1.3.1.1 Taxonomie a morfologie klíštěte obecného

V České republice patří mezi nejznámější a nejrozšířenější zástupce rodu klíšťat klíště obecné (*Ixodes ricinus*). Taxonomicky se řadí do kmene členovců (*Arthropoda*), třídy pavoukovců (*Arachnida*), řádu klíšťatovců (*Ixodida*), čeledi klíšťatovitých (*Ixodidae*), rodu klíšťat (*Ixodes*)[195].

Ploché tělo klíštěte obecného se dělí na dvě části: hlavovou (gnathosoma) a tělovou (idiosoma). Na hlavové části se nachází sací ústrojí klíštěte, na tělové části má klíště typickou hřbetní destičku (scutum), která u samců pokrývá celé tělo, zatímco u samic a nedospělých jedinců pouze jeho menší část. Pod hřbetní destičkou mají samičky klíšťat umístěný měkký zadeček, který se může několikanásobně zvětšit během sání krve na hostiteli[196]. Vnitřní stavba těla klíštěte obecného je znázorněna na obr. 11 na str. 26.



Obr. 11 Schéma vnitřní stavby těla klíštěte obecného (*Ixodes ricinu*). Převzato a upraveno podle[197].

1.3.1.2 Životní cyklus klíšťat

Klíště obecné má tříhostitelský životní cyklus. Z vajíček se v půdě vylíhnou larvy, které se přisají na prvního hostitele, nejčastěji drobné hlodavce, ještěrky nebo ptáky, ale může to být i člověk, a sají 2-6 dní. Nasátá larva odpadne ze svého hostitele a během morfologické přeměny se promění na nymfu, která se živí krví dalšího hostitele, například psa, kočky, ježka nebo člověka. Po 2 až 7 dnech sání nymfa odpadá ze svého hostitele a po další morfologické přeměně se z ní stane dospělý jedinec. Dospělí samci krev nesají, živí

se rostlinnou mízou, na hostiteli pouze hledají samici, aby se s ní spářili. Samice sají na větších zvířatech (dobytek, srnčí, psi, kočky, lišky, člověk) po dobu 1-2 týdnů. Krev je nezbytným zdrojem energie a živin pro dospělou samici, bez ní vajíčka nedozrají[196]. Schematické znázornění životního cyklu klíštěte je na obr. 12 na str. 27.

Obr. 12 Schematické znázornění tříhostitelského životního cyklu klíštěte obecného. Z vajíček se vylíhnou larvy, které musí sát krev, nejčastěji na malých zvířatech, před morfologickou přeměnou, kdy se z nich stávají nymfy. Po nasátí krve ze středně velkých zvířat prochází nymfy morfologickou přeměnou na dospělce. Dospělé samičky musí sát krev, nejčastěji na středně velkých zvířatech nebo lidech, aby byly schopné naklást snůšku vajec. Převzato a upraveno podle[198].

Ve všech fázích životního cyklu se může klíště nakazit infekčním onemocněním od hostitele a stát se tak přenašečem daného onemocnění. V České republice klíšťata nejčastěji přenášejí klíšťovou encefalitidu a lymskou boreliózu[196].

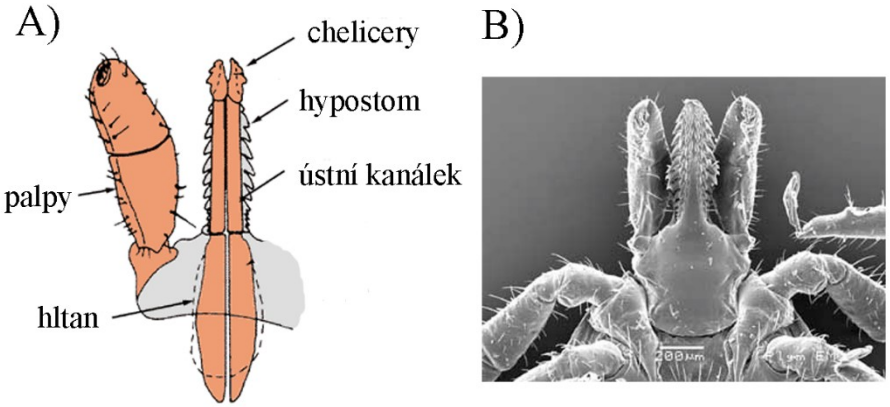
1.3.2 Trávicí systém a trávicí proteasy klíšťat

Klíšťata jsou obligatorní ektoparazité, sání krve z hostitele je nezbytné pro většinu vývojových stadií. Sání krve probíhá na povrchu hostitele, stavba těla klíštěte je k tomu přizpůsobená.

1.3.2.1 Trávicí systém klíštěte

Sací ústrojí klíštěte obecného je tvořeno hypostomem, který má tvar harpuny s dozadu zatočenými zoubky pro lepší uchycení. Hypostom je obklopen chelicery a palpy

(první a druhý pár příústních končetin). Při sání klíště odklopí palpy a vysune chelicery, kterými se přichytí ke kůži, kterou natrhne. Následně začne zasouvat chelicery a vtlačovat do ranky hypostom, kterým saje krev[199,200]. Sací ústrojí klíštěte je znázorněno na obr. 13 na str. 28.



Obr. 13 Sací ústrojí klíštěte obecného. A) Schematické znázornění sacího ústrojí. B) Fotografie z elektronového mikroskopu ventrální strany sacího ústrojí dospělé samice. Převzato a upraveno podle [201,202].

Klíšťata aktivně sají krev z hostitele za pomocí svěracích a roztahovacích svalů hltanu. Krev proudí přes hypostom ústním kanálkem, dále jícnem, který je obklopený synganglionem (centrální nervový systém), až do středního střeva, kde probíhá trávení krve. Střevo se skládá z velké hlavní komory a několika slepých výběžků (caecum), do kterých se krev dostává pomocí svalových vláken, které střevo obklopují. Kontrakce svalových buněk vedou k pohybu kapaliny v rozvětveném střevu[203]. Schéma vnitřní stavby těla je znázorněno na obr. 11 na str. 26.

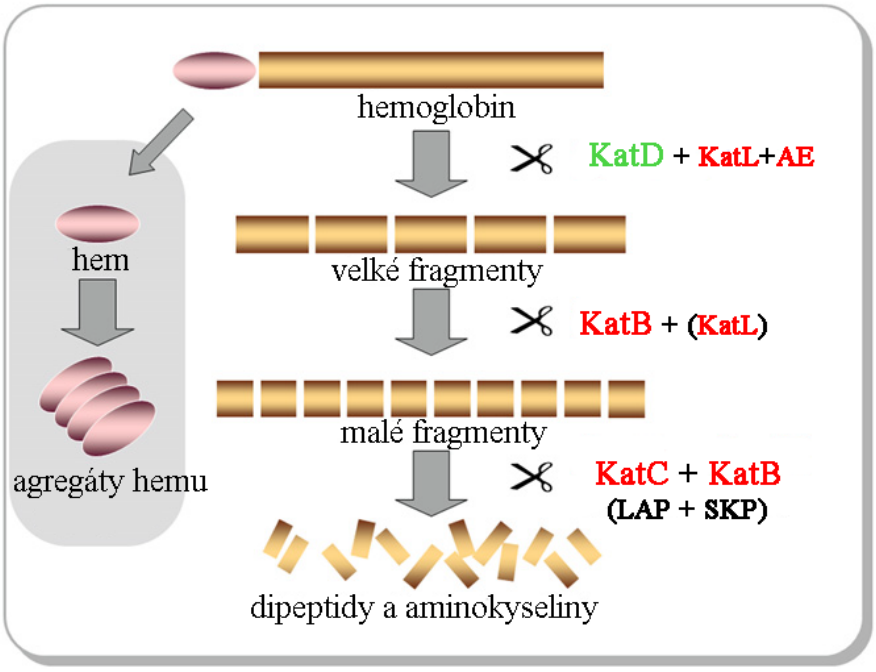
V každé fázi životního cyklu (larva, nymfa, dospělec) sají klíšťata právě jednou. Po přisátí na hostitele u dospělých samic probíhá nejprve pomalé sání, které trvá 6-9 dní, následuje rychlé sání po dobu 24-48 hodin před odpadnutím z hostitele. Celkem je samice schopná nasát až 1 ml krve, tedy stonásobek své původní hmotnosti[204].

Po nasátí krve do lumen střeva klíštěte dochází nejprve k hemolýze, tj. k popraskání membrán červených krvinek a uvolnění hemoglobinu. Následně jsou uvolněné krevní proteiny (především sérový albumin a hemoglobin) aktivně transportovány do trávicích buněk buď nespecifickou endocytózou v malých kyselých váčcích (např. sérový albumin), nebo pomocí specifických buněčných receptorů jako např. hemoglobin, který putuje do velkých endo/lysozomů, kde je dále proteolyticky štěpen[204].

1.3.2.2 Trávicí proteasy klíštěte

Proteolytická aktivita klíštěte, závislá na fázi sání klíštěte, byla popsaná při studiu degradace hemoglobinu (hemoglobinolýze) a sérového albuminu (albuminolýze)[205,206]. Fáze sání lze rozdělit na dvě části: pomalé sání (od přisátí do 6. dne) a rychlé sání (6.–8. den), po kterém následuje odpadnutí klíštěte z hostitele[205]. V prvních čtyřech dnech po přisátí je hemoglobinolytická aktivita ve střevech klíštěte velmi nízká. V následujících dvou dnech je patrný rychlý, až desetinásobný nárůst hemoglobinolýzy. V té době se ve trávicích buňkách objevuje první hemoglobin a speciální organely hemosomy, ve kterých se ukládá krystalizovaný hem z hemoglobinu. Šestý den po přisátí dosahuje hemoglobinolytická aktivita asi 65% z celkového maxima. Klíště zkonzumuje asi 2/3 z celkového objemu krve 24–48 hodin před odpadnutím[207]. Nárůst proteolytické aktivity trávicích enzymů ve střevě dále stoupá až do odpadnutí klíštěte osmý den po přisátí, kdy je nejvyšší[205].

Hemoglobinolýza probíhá intracelulárně uvnitř trávicích buněk jako kaskádovitý proces, kde nejprve působením aspartátové proteasy typu KatD vznikají velké fragmenty[207]. Tohoto kroku se v menší míře účastní i cysteinové proteasy katepsin L a legumain. Vzniklé velké fragmenty hemoglobinu jsou následně štěpeny cysteinovými proteasami katepsinem B a katepsinem L na menší fragmenty, které jsou dále degradovány aminopeptidasami a karboxypeptidasami (katepsiny B a C, leucin aminopeptidasou a serinovou karboxypeptidasou) na dipeptidy a jednotlivé aminokyseliny. Katepsiny B a C jsou v buňce asi desetkrát více zastoupeny než enzymy prvního kroku štěpení hemoglobinu[205]. Na obr. 14 na str. 30 je schematicky znázorněn proces štěpení hemoglobinu ve střevu klíštěte.



Obr. 14 Schematické znázornění kaskádovitého štěpení hemoglobinu ve střevě klíštěte obecného. Hemoglobin je v prvním kroku štěpen aspartátovou proteasou KatD, v menší míře cysteinovými proteasami katepsinem L (KatL) a legumainem (AE) na velké fragmenty, které jsou v dalším kroku štěpeny katepsinem B (KatB) a katepsinem L (KatL) na malé fragmenty. Nakonec štěpením malých fragmentů katepsinem C (KatC), katepsinem B (KatB) spolu s leucin aminopeptidasami (LAP) a serinovými karboxypeptidasami (SKP) vznikají dipeptidy a jednotlivé aminokyseliny. Hem, který se uvolní z hemoglobinu, tvoří agregáty ve specializovaných organelách hemosomech. Aspartátové proteasy jsou znázorněny zeleně, cysteinové proteasy červeně, serinové a metaloproteasy černě. Převzato a upraveno podle[205].

1.3.3 Katepsin D1 z klíštěte obecného *Ixodes ricinus*

U klíšťat bylo nalezeno několik různých aspartátových proteas, které se zapojují do řady fyziologicky důležitých procesů. Aspartátové proteasy BYC („Boophilus yolk procathepsin) a THAP („tick heme-binding aspartic protease“) se během embryogeneze podílejí na štěpení vitelinu, zásobního proteinu ve vajíčkách klíšťat[208,209]. Aspartátové proteasy typu katepsin D se významně podílejí na trávení proteinů z krve, především hemoglobinu a sérového albuminu, důležitých zdrojů aminokyselin a energie pro dospělé samice[205,210]. U klíšťat rodu *Ixodes ricinus* byly popsány tři proteasy typu katepsin D (IrKatD1–3), které se vyskytují v různých vývojových stadiích a v různých tkáních[72]. IrKatD1 byl nalezen ve střevech sajících klíšťat. Ve střevech a ve slinách klíštěte byl detekován IrKatD2. IrKatD3 byl detekován především v ovariích[211].

Všechny tři izoenzymy proenzymu IrKatD mají podobnou primární strukturu. Liší se pouze v propeptidové části, která u neaktivního proenzymu blokuje aktivní centrum enzymu. Rozdíly jsou v délce i složení aminokyselinové sekvence propeptidu. Zatímco N-

koncová část propeptidu je konzervovaná, jeho C-koncová sekvence je variabilní. IrKatD3 má nejdelší propeptid (37 aminokyselin), IrKatD2 má kratší propeptid (28 aminokyselin) a IrKatD1 má propeptid dlouhý pouze 22 aminokyselin. Všechny izoenzymy mají dva katalytické aspartáty (Asp33 a Asp231) a polyprolinovou smyčku[211]. Předpokládá se, že IrKatD1 a IrKatD3 mají dvě glykosylační místa, IrKatD2 má jedno glykosylační místo navíc[211].

1.3.3.1 Fyziologie katepsinu D1 z klíštěte obecného

mRNA z IrKatD1 byla nalezena pouze ve střevech sajících klíšťat a její množství přibývalo v čase po dobu sání, nejvíce ho bylo šestý den. Podobný nárůst byl pozorován s množstvím proteinu IrKatD1 ve střevech[211].

Sérový albumin a hemoglobin z krve hostitele jsou primárním zdrojem živin pro klíště[204]. IrKatD1 v kyselém prostředí endo/lysozomů zahajuje štěpení hemoglobinu jako dominantní proteasa, ale jeho podíl na štěpení sérového albuminu je oproti štěpení hemoglobinu významně nižší (cca 10%)[205]. Podobně jako u hemoglobinu je samotný IrKatD1 schopný štěpit sérový albumin pouze na velké fragmenty[206].

1.3.3.2 Biochemická charakterizace katepsinu D1 z klíštěte obecného

Substrátová specifita IrKatD1, která byla zkoumána pomocí knihovny syntetických peptidových substrátů a přirozeného substrátu hemoglobinu, je podobná jako u ostatních aspartátových proteas typu KatD. IrKatD1 preferuje štěpení mezi hydrofobními aminokyselinami. Do podmísta S_1 se nejlépe váže Phe, Tyr, Leu, Trp nebo Nleu, nikoliv Ala, Val nebo Ile. Do S_1' podmísta se nejlépe vážou také hydrofobní aminokyseliny až na Pro a Leu. Vazba do ostatních podmíst není tak specifická, až na preferenci Tyr v pozici P_3 a Ala v pozici P_2'[211].

IrKatD1 je aktivní podobně jako jiné aspartátové proteasy v kyselém pH v rozmezí pH 2,5 a pH 5,0. Jeho aktivita je účinně blokovaná obecným inhibitorem aspartátových proteas pepstatinem a proteinovým inhibitorem PDI z lilku brambor. Syntetický inhibitor aspartátové HIV proteasy lopinavir inhibuje IrKatD1 pouze částečně. Proenzym IrKatD1 (proIrKatD1) je schopen se v kyselém pH jednokrokově aktivovat, k odštěpení propeptidu dochází mezi zbytky Lys22p-Ile1[211].

2 Cíle disertační práce

Disertační práce se zabývá katepsinem D (KatD), který hraje důležitou roli v řadě fyziologicky významných proteolytických procesů. Ačkoliv nedostatečná regulace KatD je spojena s řadou patologií, doposud nebyly popsány žádné endogenní inhibitory KatD u vyšších organismů. Cílem práce je identifikace a charakterizace těchto endogenních inhibičních molekul, analýza mechanismů jejich působení a navrhování nových biomimetických regulátorů KatD.

Dílčí cíle disertační práce jsou:

1. Popsat působení bioaktivních sfingolipidů na funkci lidského KatD. Zejména analyzovat vztah mezi strukturou sfingolipidů a proteolytickou aktivitou KatD a vliv faktorů prostředí na interakci. Diskutovat možný vztah k patofyziologii.
2. Na katepsinu D z klíštěte obecného (IrKatD1) analyzovat strukturně funkční vztahy v tzv. aktivačním peptidu (propeptidu) s využitím proteinové krystalografie a biochemických metod. Identifikovat inhibiční motiv ve struktuře propeptidu, studovat jeho interakci s IrKatD1 a diskutovat význam tohoto regulačního mechanismu pro aspartátové proteasy.
3. Studovat funkční vlastnosti biomimetických makrocyklických inhibitorů odvozených ze struktury pepstatinu, mikrobiálního inhibitoru aspartátových proteas, zejména inhibiční interakci s lidským KatD, selektivitu a vlastnosti umožňující posoudit jejich potenciální využití pro vývoj léčiv.

3 Materiály a metody

Tato kapitola stručně popisuje základní materiály, vybavení a metodiky použité během vypracování disertační práce. Podrobné informace jsou uvedené v přiložených publikacích.

3.1 Materiály a laboratorní vybavení

Většina výsledků byla získaná s využitím laboratorního vybavení na Ústavu organické chemie a biologie AV ČR (ÚOCHB). Sběr difrakčních dat pro rentgenostrukturní analýzu proteinových krystalů proběhl na synchrotronu Bessy II electron storage ring v Helmholtz-Zentrum v Berlíně v Německu a na synchrotronu ESRF v Grenoblu ve Francii. Krystalové struktury byly řešeny ve spolupráci s laboratoří Strukturní biologie na ÚOCHB. Peptidové a peptidomimetické inhibitory a nekomerční substráty byly syntetizovány ve skupině Medicinální chemie na ÚOCHB. Sady komerčních krystalizačních roztoků byly od firem Jena Bioscience, Molecular Dimensions a Hampton Research.

3.2 Metody

Uveden je výčet hlavních metod popsaných v přiložených publikacích:

Metody molekulární biologie:

Klonování genu proIrKatD1 do plasmidu pET101/D-TOPO, mutageneze katalytického aspartátu D231N proIrKatD1 konstruktu, transformace buněk *E. coli* BL21(DE), rekombinantní exprese dvou forem proIrKatD1 (přirozené formy a katalyticky neaktivní formy) v *E. coli*.

Biochemické metody:

Izolace proIrKatD1 z inkluzních tělísek, renaturace proIrKatD1, jeho aktivace a purifikace chromatografickými metodami (ionexová, gelová a afinitní chromatografie); chromatografická izolace lidského KatD z tkání placent. Elektroforetická separace proteinů na SDS-PAGE, separace peptidů na RP-HPLC, určení N-koncové sekvence proteinů přenesených na polyvinyldifluoridovou membránu, kvantifikace peptidových substrátů a inhibitorů pomocí aminokyselinové analýzy.

Enzymologické metody:

Měření aktivit enzymů na fluorescenční čtečce (Tecan) pomocí FRET (Fluorescence Resonance Energy Transfer) peptidových substrátů, stanovení kinetických parametrů IC_{50},

K_m, K_i, k_{cat}, určení módu inhibice. Stanovení hodnoty disociační konstanty K_d s využitím měření termoforetického efektu na přístroji Monolith NT.115 (NanoTemper).

Krystalografické metody:

Příprava molekulárních forem a inhibičních komplexů IrKatD1 pro krystalizaci, vyhledání krystalizačních podmínek a jejich optimalizace. Analýza intermolekulárních interakcí ve 3D modelech a molekulární grafika.

4 Výsledky

Výsledky disertační práce jsou prezentovány ve třech článcích publikovaných v mezinárodních recenzovaných časopisech. Kapitoly 4.1 až 4.3 obsahují komentář s hlavními výsledky článků a uvedením podílu autora disertační práce na publikacích. Články se zabývají přirozenými mechanismy inhibice KatD a jsou zaměřené na nové endogenní inhibitory lidského KatD (publikace č. 1) a modelového parazitárního KatD (publikace č. 2) a dále na vývoj biomimetických inhibitorů lidského KatD (publikace č. 3).

Seznam publikací:

Publikace č. 1: Complex modulation of peptidolytic activity of cathepsin D by sphingolipids.

Žebrakovská I., Máša M., Srp J., Horn M., Vávrová K., Mareš M.

Biochim Biophys Acta, 12, 1097-1104 (2011). (IF = 4,4)

Publikace č. 2: Novel structural mechanism of allosteric regulation of aspartic peptidases via an evolutionarily conserved exosite.

Hánová I., Brynda J., Houštecká R., Alam N., Sojka D., Kopáček P., Marešová L., Vondrášek J., Horn M., Schueller-Furman O., Mareš M.

Cell Chem Biol, 25, 318-329 (2018). (IF = 6,8)

Publikace č. 3: Biomimetic macrocyclic inhibition of human cathepsin D: structure-activity relationship and binding mode analysis.

Houštecká R., Hadzima M., Fanfrlík J., Brynda J., Pallová L., **Hánová I.**, Mertlíková-Kaiserová H., Lepšík M., Horn M., Smrčina M., Majer P., Mareš M.

J Med Chem, 63, 1576-1596 (2020). (IF = 6,1)

4.1 Publikace č.1: Complex modulation of peptidolytic activity of cathepsin D by sphingolipids

4.1.1 Souhrn

Působení bioaktivních sfingolipidů lipidů na aktivitu proteolytických enzymů nebylo doposud detailně studováno. Tato studie přináší první enzymologickou analýzu funkční regulace KatD přímou interakcí s těmito lipidickými molekulami. Nové informace významně rozšiřují současný pohled na komplexní vztahy mezi KatD a proteiny účastnící se metabolismu sfingolipidů[116].

Pomocí panelu 38 přirozených i syntetických derivátů sfingolipidů byl detailně studován vztah mezi strukturou sfingolipidu a působením na aktivitu KatD. Určité sfingosiny a ceramidy byly účinnými reversibilními inhibitory KatD vykazující až sub-mikromolárních hodnoty inhibičních parametrů, pro nejlepší inhibitor N-oktyl-sfingosin byla naměřena hodnota IC_{50} = 0,14 μM. Pro fosforylované deriváty sfingolipidů typu sfingosin-1-fosfátu a ceramid-1-fosfátu byla nalezena naopak pozitivní modulace KatD vedoucí až k několikanásobnému nárůstu aktivity. Tento efekt byl způsoben zvýšením afinity enzymu k substrátu (K_m), což vedlo k vyšší katalytické účinnosti KatD (k_{cat}/K_m). Modulace aktivity KatD působením sfingolipidů byla obecně závislá na pH s nejvýraznějšími projevy inhibice, resp. aktivace v mírně kyselém pH. V přítomnosti fosfosfingolipidů docházelo k výraznému posunu pH optima pro lidský KatD z pH cca 4,0 na pH 5,0.

Mechanismus modulace aktivity KatD pomocí biogenních sfingolipidů byl studován několika přístupy a byla prokázána jejich přímá interakce s KatD. Disociační konstanta inhibitorů byla určena pomocí měření polarizace fluorescence. Kompetitivní mód inhibice ukazoval na alespoň částečnou vazbu sfingolipidů do aktivního centra enzymu, což je v souladu se substrátovou specifitou KatD, kde jsou v kritických pozicích preferovány hydrofobní aminokyseliny. Nicméně určení přesného vazebného místa pro sfingolipidy a fosfosfingolipidy bude vyžadovat použití metod strukturní analýzy.

Dále byl testován vliv vybraných faktorů a liposomů, které v experimentu simulovaly fyziologické prostředí a biologické membrány, na interakci KatD se sfingolipidy. Získané výsledky ukazují, že modulace aktivity KatD biogenními sfingolipidy je relevantní pro prostředí buněk a tkání. Bylo také prokázáno, že tato modulace je vysoce specifická pro KatD, protože nebyl pozorován vliv studovaných

sfingolipidů na modelové zástupce proteas z jiných tříd, včetně aspartátových proteas, kam KatD patří.

Můj podíl na práci zahrnoval: (1) izolaci lidského KatD, (2) měření aktivity KatD ve fluorescenčním kinetickém testu v přítomnosti sfingolipidů, určení inhibičních konstant IC_{50} a módu inhibice, (3) měření aktivit zástupců různých tříd proteas v přítomnosti sfingolipidů, (4) analýzy alternativního aktivitního testu s peptidovým substrátem KatD pomocí HPLC, (5) přípravu manuskriptu.

4.1.2 Publikace č. 1

Biochimica et Biophysica Acta 1811 (2011) 1097–1104

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbalip

Complex modulation of peptidolytic activity of cathepsin D by sphingolipids

Iva Žebrakovská [a,b], Martin Máša [a], Jaroslav Srp [a,b], Martin Horn [a], Kateřina Vávrová [c], Michael Mareš [a,*]

[a] Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nam. 2, 16610 Prague, Czech Republic
[b] Department of Biochemistry, Faculty of Science, Charles University, Albertov 6, 12843 Prague, Czech Republic
[c] Department of Inorganic and Organic Chemistry, Faculty of Pharmacy, Charles University, Heyrovskeho 1203, 50005 Hradec Kralove, Czech Republic

ARTICLE INFO

Article history:
Received 5 May 2011
Received in revised form 22 August 2011
Accepted 12 September 2011
Available online 23 September 2011

Keywords:
Sphingolipid
Phospholipid
Inhibition
Activation
Cathepsin D
Enzyme regulation

ABSTRACT

Cathepsin D is an aspartic peptidase involved in cellular processes including proliferation and apoptosis and implicated in human pathologies such as cancer and neurodegeneration. Our knowledge about the relationship between proteolysis and bioactive sphingolipids is still very limited. Here, we describe a complex pattern of modulation of the peptidolytic activity of cathepsin D by sphingolipids. A panel of sphingolipid derivatives was screened in a FRET-based assay; these molecules demonstrated negative or positive modulation of cathepsin D peptidolytic activity, depending on the sphingolipid structure. Certain sphingosines and ceramides inhibited cathepsin D in the submicromolar range, and structural requirements for this inhibitory effect were evaluated. The interaction of cathepsin D with sphingolipids was also demonstrated by fluorescence polarization measurements and determined to follow a competitive inhibition mode. In contrast, monoester phosphosphingolipids, especially ceramide-1-phosphate, were identified as activators of cathepsin D peptidolytic activity at submicromolar concentrations. Thus, sphingolipids and phosphosphingolipids, known to be antagonistic in their cell-signaling functions, displayed opposite modulation of cathepsin D. Sphingolipid-based modulators of cathepsin D are potentially involved in the control of cathepsin D-dependent processes and might serve as a scaffold for the development of novel regulators of this therapeutic target.

1. Introduction

Sphingolipids (SLs) are potent signaling molecules capable of regulating vital cellular functions including cell growth, differentiation, senescence, apoptosis, and intracellular trafficking [1,2]. SLs are critically involved in tumor development and cancer cell death; *e.g.* ceramide has been recognized as a "tumor-suppressor lipid" based on its ability to block tumor initiation and metastasis [3,4]. Thus, SLs and their analogs represent a promising tool for manipulating the SL network in anticancer therapy. Despite recent advances in our understanding of SL metabolism and their roles in modulating cell responses, much less is known about SL-protein interactions and, in particular, about how SLs modulate the biochemical activity of target enzymes [5]. Our work focuses on cathepsin D (CatD), the only proteolytic enzyme that directly binds SLs reported to date [6].

CatD is an aspartic peptidase of general importance for cell growth and tissue homeostasis, as indicated by gene knockout studies [7]. It is localized in lysosomes and participates in the turnover of cellular proteins as well as in the selective processing of antigens, hormones, and growth factors. In cancer cells, CatD is overexpressed and secreted in the tumor microenvironment; it is associated with cancer development and progression through its complex proliferative and proteolytic action [8,9]. CatD is also an independent marker of poor prognosis in breast cancer [10,11]. During apoptosis, CatD is released into the cytosol, where it can interact with pro-apoptotic, anti-apoptotic, or nuclear proteins [8,9,12].

Evidence accumulated in the past decade indicates that there are multiple links between the physiological actions of CatD and bioactive SLs. Intracellular CatD zymogen interacts with prosaposin, a precursor of the sphingolipid activator proteins (saposins A-D), and this interaction supports CatD targeting and activation as well as prosaposin processing [13]. Extracellular CatD zymogen can be directed to lipid rafts [14], cell membrane structures that are generally regulated by SLs [15], by the CatD-binding protein LRP1. There is a functional link between production of intracellular ceramide by endolysosomal acid sphingomyelinase and CatD maturation [6,16]. Ceramide levels have been shown to regulate CatD maturation and transport in apoptosis models [16,17]. Heinrich et al. reported that SLs can induce CatD-mediated proteolysis [6,16]. Importantly, experiments using SL-affinity chromatography and SL-based photoreactive probes determined that CatD directly binds SLs such as sphingosine and ceramide [6].

In the present work, we performed the first detailed enzymological analysis of the interaction between SLs and CatD with regard to its impact on the peptidolytic activity of CatD. We show that specific SLs and

Abbreviations: CatD, cathepsin D; FRET, fluorescence resonance energy transfer; PSL, phosphosphingolipid; SL, sphingolipid

* Corresponding author at: Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nam. 2, 16610 Prague 6, Czech Republic. Tel.: +420 220183358; fax: +420 220183578.
E-mail addresses: zebrakovska@uochb.cas.cz (I. Žebrakovská), mosha@atlas.cz (M. Máša), srp@uochb.cas.cz (J. Srp), horn@uochb.cas.cz (M. Horn), katerina.vavrova@faf.cuni.cz (K. Vávrová), mares@uochb.cas.cz (M. Mareš).

1388-1981/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbalip.2011.09.005

phosphosphingolipids (PSLs) are complex modulators of CatD capable of inhibiting and activating CatD, respectively. No endogenous mammalian CatD inhibitors are known and SLs therefore represent unique inhibitory regulators of this enzyme. We suppose that the SL-PSL network modulating CatD might be relevant in a physiological context.

2. Materials and methods

2.1. Materials

Human CatD was purified to homogeneity from placenta [18]. Pepsin and papain were from Sigma-Aldrich (MO, USA), and trypsin was from Serva (Germany). Specific peptidic substrates for CatD, namely the FRET (fluorescence resonance energy transfer)-based substrate Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu (Abz,1-aminobenzoic acid; Nph, 4-nitrophenylalanine) and the CatD propeptide-derived substrate [15p-36p], were designed and synthesized as previously described [19]. Z-Phe-Arg-AMC substrate (AMC, 7-amino-4-methylcoumarin) was from Bachem (Switzerland). Lipids (see compound numbers in Tables 1 and S1) were obtained from the following sources: Sigma-Aldrich: **1**, **3**, **4**, **6**, **8**, **12**, **13**, **16**–**19**, **21**, **26**–**29**, **36**–**38**; Cayman Chemical Company (MI, USA): **2**, **11**, **33**–**35**; Avanti Polar Lipids (AL, USA): **10**, **22**; Invitrogen (CA, USA): **20**; and Matreya (PA, USA): L-*threo* stereoisomer of **12**. Compounds **5**, **7**, **9**, **14**, **15**, **23**–**25**, **30**–**32**, and L-*threo*-stereoisomer of **1** were prepared as previously described [20–22]. Heparin (18 kDa) and pepstatin A were from Sigma-Aldrich. Unilamellar liposomes (100 nm in diameter) were prepared from 1,2-dimyristoyl-phosphocholine (Avanti Polar Lipids) in water according to the manufacturer's protocol.

2.2. FRET-based assay of CatD activity

The peptidolytic activity of CatD and its modulation by lipids was assayed with the FRET-based optimized CatD substrate Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu [19]. The reaction was performed in microplate format in a total assay volume of 100 µl. The reaction mixture typically contained 2 nM CatD and 0–200 µM lipid in 0.1 M sodium acetate, pH 4.2 (or pH 3.0–6.0 where indicated), containing 0.1% polyethylene glycol 1500. Where indicated, heparin or salts were added to the reaction mixture to give a final concentration of 750 µg heparin/ml or 1–250 mM NaCl, $CaCl_2$, or $MgCl_2$. In liposome experiments, 100 nm liposomes were added to the reaction mixture to give a final concentration of 100 µM. The mixture was incubated at 37 °C for 10 min, followed by the addition of substrate to a final concentration of 20 µM. The kinetics of product release were continuously monitored in an Infinite M200 microplate reader (Tecan) at excitation and emission wavelengths of 330 and 410 nm, respectively. Lipid stock solutions (10 mM) were prepared in dimethyl sulfoxide (alternatively in acetonitrile, methanol, or ethanol); the solvent concentration in the assay was 2%. IC_{50} values were determined from the residual velocities using the dose–response plot (v_i/v_0 vs [I]) analyzed by Grafit (Erithacus software, UK). The kinetic parameters (K_m and k_{cat}) and inhibition mode were determined using the same activity assay with substrate concentrations up to 400 µM. The initial velocities of product release were analyzed by non-linear regression (using Grafit) and by a Lineweaver-Burk plot. The reversibility of CatD inhibition by sphingosine was tested in a dilution experiment. CatD (20 nM) was preincubated (10 min at 37 °C) in 0.1 M sodium acetate, pH 4.2, with 0.4 or 4 µM sphingosine (corresponding to ~IC_{50} and 10 × IC_{50}, respectively). The mixture was then diluted 10-fold with the same buffer containing 0.4 µM sphingosine (for IC_{50} preincubation concentration) or without inhibitor (for 10 × IC_{50} preincubation concentration) and incubated for 10 min at 37 °C. The final activity was then measured with the FRET substrate as described above. All activity measurements were performed in triplicate.

Table 1
Modulation effects of sphingolipids and related compounds on the peptidolytic activity of CatD[a].

Comp. no.	Abbreviation	Inhibition	Activation
		IC_{50} (µM) ± SE	Fold increase ± SE
	Sphingosine derivatives		
1	Sphingosine	0.39 ± 0.04	n.e.
2	*N,N*-dimethylsphingosine	0.45 ± 0.11	n.e.
3	*N,N,N*-trimethylsphingosine	0.16 ± 0.01	n.e.
4	Azidosphingosine	n.e.	n.e.
5	C12-sphingosine	>200	n.e.
6	Dihydrosphingosine (sphinganine)	0.41 ± 0.03	n.e.
7	C20-sphingosine	0.22 ± 0.02	n.e.
8	Sphingosine-1-phosphate	n.e.	2.0 ± 0.1
9	(*2S,3R*)-2-aminoicos-4-yne-1,3-diol (triple bond C20-sphingosine)	0.27 ± 0.03	n.e.
10	Phytosphingosine	0.30 ± 0.04	n.e.
11	C8-ceramine (*N*-octylsphingosine)	0.14 ± 0.03	n.e.
	Ceramide derivatives		
12	C6-ceramide (C18 base)	1.3 ± 0.1	n.e.
13	C6-dihydroceramide (C18 base)	n.e.	n.e.
14	C2-ceramide (C12 base)	n.e.	n.e.
15	C6-ceramide (C12 base)	69.1 ± 8.5	n.e.
16	C2-ceramide (C18 base)	0.31 ± 0.03	n.e.
17	C8-ceramide-1-phosphate	n.e.	5.0 ± 0.1
18	1-*O*- β-D-galactosyl-ceramide	61.7 ± 8.1	n.e.
19	Sphingomyelin	n.e.	n.e.
20	BODIPY FL C5-ceramide	11.7 ± 1.9	n.e.
21	C18-ceramide-1-phosphate	n.e.	3.4 ± 0.9
22	C6-pyridinium-ceramide	0.41 ± 0.06	n.e.
23	C24-NBD-ceramide	n.e.	n.e.
24	C12-NBD-ceramide	4.3 ± 0.7	n.e.
25	C6-NBD-ceramide	6.9 ± 0.7	n.e.
26	3-*O*-benzoyl-C16-ceramide	n.e.	n.e.
	Others		
27	1,2-dipalmitoyl-glycerol	n.e.	n.e.
28	(*1R,2R*)-2-amino-1-phenyl-1,3-propanediol	n.e.	n.e.
29	(*1S,2S*)-2-amino-1-phenyl-1,3-propanediol	n.e.	n.e.
30	*N*-dodecanoylglycine dodecyl ester	n.e.	n.e.
31	*N*-dodecanoyl-L-serine dodecyl ester	n.e.	n.e.
32	*N*-oleoyl-L-serine dodecyl ester	n.e.	n.e.
33	Prostaglandin E1	n.e.	n.e.
34	Myriocin	n.e.	n.e.
35	(*1R,2R*)-2-myristylamino-1-(4-nitrophenyl)-1,3-propandiol	n.e.	n.e.
36	Phosphatidic acid	n.e.	4.9 ± 0.2
37	Phosphatidylcholine	n.e.	n.e.
38	Phosphatidylethanolamine	n.e.	n.e.

[a] CatD activity was measured with a FRET-based octapeptide substrate at pH 4.2 in the presence of the tested compounds. Structures of the compounds are given in Supplemental Table S1. The inhibitory effect was evaluated by determining the IC_{50} value. The activation effect was expressed as relative fold increase of CatD activity induced by 10 µM compound compared to the non-activated control. n.e. denotes no modulation effect with 200 µM compound. Mean values ± SE are given.

2.3. Activity assays for other proteases

The peptidolytic activities of trypsin and papain were assayed with the fluorogenic substrate Z-Phe-Arg-AMC. The reaction mixture contained 60 pM trypsin or 10 pM papain in 0.1 M Tris–HCl, pH 7.0, containing 0.1% polyethylene glycol 1500, or 0.1 M bis-Tris–HCl, pH 6.0, containing 2.5 mM dithiothreitol and 0.1% polyethylene glycol 1500, respectively. Where indicated, lipids were added to the reaction mixture to give a final concentration of 1 µM or 10 µM. The mixture was incubated at 37 °C for 10 min, followed by the addition of

50 μM Z-Phe-Arg-AMC. The instrumentation and measurement were as detailed in Section 2.2, except that the excitation and emission wavelengths were 360 and 465 nm, respectively. The peptidolytic activity of pepsin was determined as described for CatD (Section 2.2) with the following modifications: the reaction mixture contained 0.2 nM pepsin and was buffered with 0.1 M sodium acetate, pH 4.0, containing 0.1% polyethylene glycol 1500.

2.4. Fragmentation of peptide [15p-36p] by CatD

The reaction mixtures, composed of 50 μM peptide [15p-36p] and 1.25 nM CatD with 0–10 μM C8-ceramide-1-phosphate or 3.75 nM CatD with 0–10 μM sphingosine, were incubated in 0.1 M sodium acetate, pH 4.2, at 37 °C for 12 h. The reactions were stopped by adding trifluoroacetic acid (TFA) to a final concentration of 0.5% (v/v). The mixtures were separated by RP-HPLC using an Agilent 1200 Series system (Agilent Technologies) on a Kromasil C18 column (250×4.6 mm, 5 μm; Alltech) equilibrated in 0.1% (v/v) TFA and eluted with a linear 1.5%/min gradient of 99% (v/v) acetonitrile in 0.1% (v/v) TFA. The collected peak fractions were analyzed by ESI mass spectrometry on an LCQ Classic Finnigan Mat (Thermo Finnigan).

2.5. Fluorescence polarization measurements

Fluorescence polarization measurements were performed with an Aminco-Bowman Series 2 luminescence spectrophotometer (Thermo Electron) at 37 °C using excitation and emission wavelengths of 505 and 514 nm, respectively. Anisotropy titration was carried out in competition experiments with 0.001–40 μM sphingosine, 2 μM BODIPY FL C5-ceramide (Invitrogen) as a fluorescent "reporter" probe, and 50 nM CatD in 0.1 M sodium acetate, pH 5.5. The binding mixture was incubated for 45 min at 37 °C to reach equilibrium. The fluorescence of the buffer alone was subtracted from each measurement, and the anisotropy was calculated according to [23]. G-factor correction was performed to eliminate instrumental bias.

The anisotropy data were converted to the fraction of the probe bound using the following equation:

$$f_b = \frac{r_{obs} - r_0}{(r_{obs} - r_0) + R(r_{\max} - r_{obs})}$$

where f_b is the fraction of the probe bound and r_{obs}, r_0, and r_{max} are the observed anisotropy, anisotropy of the free probe, and anisotropy of the probe with CatD, respectively. R is the fluorescence intensity of the probe with CatD divided by the intensity of the free probe. The K_d of the binding of sphingosine to CatD was determined by nonlinear least-squares fit to a single-site binding model using Grafit software.

2.6. Analytical methods

The stock concentration of CatD was determined by titration with pepstatin A using the FRET-based assay. The stock solutions of peptide substrates were quantified by amino acid analysis. The stock solutions of SLs were quantified according to [24]; the acylated SLs were subjected to acidic hydrolysis prior to analysis.

3. Results

3.1. Structure-activity profiling of SLs reveals complex modulation of CatD activity

A panel of SL derivatives and structurally related compounds was screened for the ability to influence the peptidolytic activity of human CatD (Tables 1 and S1). An *in vitro* FRET-based assay with the fluorogenic peptidic substrate Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu [19] was employed to measure CatD activity in the presence of SLs. The results presented in Table 1 show that these compounds display a complex pattern of modulation of CatD activity, including inhibition or activation; some derivatives had no significant effect.

The inhibitory compounds were evaluated by determining their IC_{50} values, which generally ranged from submicromolar to micromolar values (Table 1). D-*erythro*-sphingosine and C6-ceramide (compounds **1** and **12**, Table 1) were used as model SL structures; these compounds were compared with other derivatives to assess the structural requirements for inhibitory interaction. The naturally occurring D-*erythro*-sphingosine with C18 chain length (**1**) inhibited CatD with an IC_{50} of 0.39 μM. All the other sphingoid bases, including sphingosines with various chain lengths (**5**, **7**), *N*-alkylsphingosines (**2**, **3**, **11**), dihydrosphingosine (**6**), phytosphingosine (**10**), and a triple bond-containing analog (**9**) were able to inhibit CatD to different extents, with the only exception being an azidosphingosine analog (**4**). Thus, the inhibition was independent of the presence of the C4 *trans*-double bond (compare **1**, **6**, **10,** and **9**). The gradual shortening of the sphingosine chain length from C20 (**7**) to C12 (**5**) dramatically reduced the CatD inhibition (IC_{50} from 0.22 to >200 μM). The importance of the presence of an alkyl chain of a certain length was further confirmed by the lack of activity of **28** and **29**, which possess a phenyl ring instead of the pentadecenyl chain.

In the ceramide panel, C6-ceramide (with natural C18 sphingosine base, **12**) inhibited CatD with an IC_{50} of 1.4 μM, and the effect was abrogated after saturation of the C4 *trans*-double bond (**13**). Thus, an intact allylic hydroxyl in ceramides is critical for inhibition of CatD. This was also confirmed by the inactivity of **26**, which has the allylic hydroxyl masked with benzoyl group, and of **31** and **32**, in which the allylic hydroxyl moiety was replaced by an ester bond. Shortening of the acyl chain to C2-ceramide (**16**) markedly increased its inhibition potency (IC_{50} = 0.31 μM). On the other hand, the same change in the ceramide with the base chain truncated to C12 (**15** and **14**) resulted in a loss of inhibitory potency. Introduction of a bulky fluorescent label at the end of the acyl chain, *e.g.* in Bodipy-C5-ceramide (**20**) or the NBD-ceramides **23**–**25**, generally decreased the inhibitory effect towards CatD. Nevertheless, a similar acyl chain modification with pyridinium salt (highly soluble compound **22**) yielded a more effective inhibitor compared to the unmodified ceramide **12**.

The stereospecificity of SL-mediated inhibition of CatD was probed using unnatural L-*threo* stereoisomers of the natural D-*erythro* structures of sphingosine (**1**) and C6-ceramide (**12**). The IC_{50} values were rather similar for the sphingosine isomers (0.48 ± 0.06 μM for L-*threo* compared with 0.39 ± 0.04 μM for D-*erythro*), but differed to a

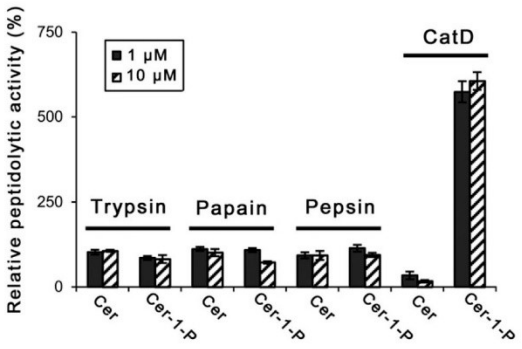


Fig. 1. Comparison of modulatory effects of sphingolipids on CatD and representative proteases. Activities of trypsin, papain, pepsin, and CatD were determined in the presence of 1 μM and 10 μM concentrations of C6-ceramide (Cer) and C8-ceramide-1-phosphate (Cer-1-P). The peptidolytic activities of pepsin and CatD were measured with a FRET-peptide substrate at pH 5.0 and 4.0, respectively, and those of papain and trypsin with a fluorogenic AMC-peptide substrate at pH 6.0 and 7.0, respectively. For each protease, mean values ± SD are expressed relative to the control experiment without lipid (100%).

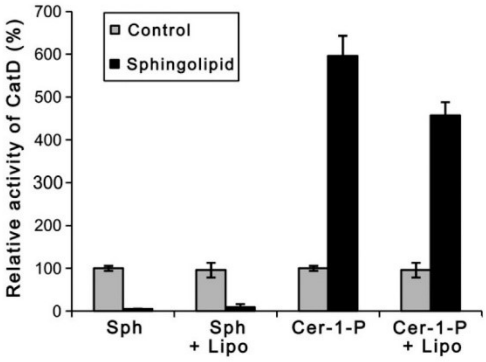


Fig. 2. Modulatory effects of sphingolipids on CatD activity in the presence of liposomes. The peptidolytic activity of CatD was measured in a FRET-based assay at pH 5.0 in the presence of 10 μM sphingosine (Sph) or 10 μM C8-ceramide-1-phosphate (Cer-1-P). The experiment was performed in the presence or absence of 100 μM liposomes (Lipo). The sphingosine-mediated inhibition and C8-ceramide-1-phosphate-mediated activation are expressed relative to the untreated control without sphingolipids and liposomes (100%). Mean values ± SD are given.

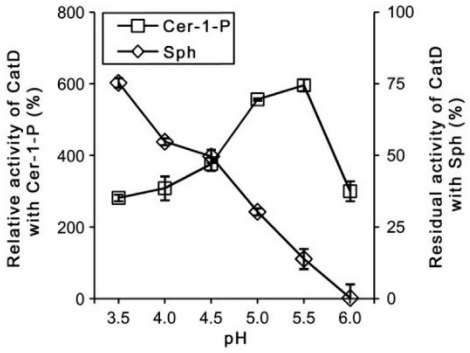


Fig. 4. Modulatory effects of sphingolipids on CatD activity are pH dependent. The peptidolytic activity of CatD was measured in a FRET-based assay in the presence of 0.4 μM sphingosine (Sph) or 1 μM C8-ceramide-1-phosphate (Cer-1-P) and compared with the untreated controls. For C8-ceramide-1-phosphate-mediated activation (left axis), the values are expressed relative to the untreated controls at the same pH (100%). For sphingosine-mediated inhibition (right axis), the values are expressed as the remaining activity relative to the untreated controls at the same pH (100%). Mean values ± SD are given.

greater extent for the C6-ceramide isomers (3.3 ± 0.3 μM for L-*threo* compared with 1.3 ± 0.1 μM for D-*erythro*).

SLs that potentiated CatD activity were evaluated by determining their activation effect at a 10 μM concentration (Table 1). A high effect was observed for short-chain C8-ceramide-1-phosphate (**17**) and naturally occurring long-chain ceramide-1-phosphate (with prevailing C18 acyl, **21**). Compounds **17** and **21** induced the activity of CatD by about 5-fold and 3.5-fold, respectively. In contrast, only ~2-fold induction was found for sphingosine-1-phosphate (**8**). The phosphate monoester group of these PSLs was essential for the activation effect; sphingomyelin **19** (with choline linked to the phosphate) was inactive, and cerebroside **18** (with the phosphate replaced with a monosaccharide) was weakly inhibitory. Interestingly, the activation effect was also found for phosphatidic acid (**36**) but not the phosphate diesters phosphatidylcholine (**37**) and phosphatidylethanolamine (**38**).

The modulation of CatD activity by SLs and PSLs was compared with the action of the lipids on other peptidases. The selected enzymes pepsin, papain, and trypsin are typical representatives of the aspartic, cysteine, and serine classes of peptidases, respectively. Their peptidolytic activities were measured in the presence of C6-ceramide (a CatD inhibitor) and C8-ceramide-1-phosphate (a CatD activator). Fig. 1 shows that these lipids do not have pronounced effects on the tested peptidases, in contrast to their modulatory effects on CatD. This demonstrates the specificity of the SL/PSL-mediated modulation of CatD activity.

Finally, we investigated the modulatory effects of SLs and PSLs in the presence of liposomes to simulate the environment of biological membranes. Fig. 2 shows that CatD was efficiently inhibited by sphingosine and was activated by ceramide-1-phosphate at a 10-fold molar excess of unilamellar liposomes over lipid. This experiment strongly suggests that CatD is capable of functional interaction with SLs/PSLs in membrane-containing systems.

3.2. Inhibitory interaction of SLs with CatD and its mechanism

The inhibitory interaction of SLs with CatD was characterized in detail. The inhibition curves of the peptidolytic activity of CatD measured with the FRET octapeptide substrate are given in Fig. 3A; sphingosine (**1**) and C6-ceramide (**12**) exerted concentration-dependent inhibition of CatD in the submicromolar and low micromolar range, respectively (Table 1). Moreover, their inhibition potency was influenced by pH. The pH profile of the most inhibitory natural SL, sphingosine, demonstrated that the extent of its CatD inhibition substantially

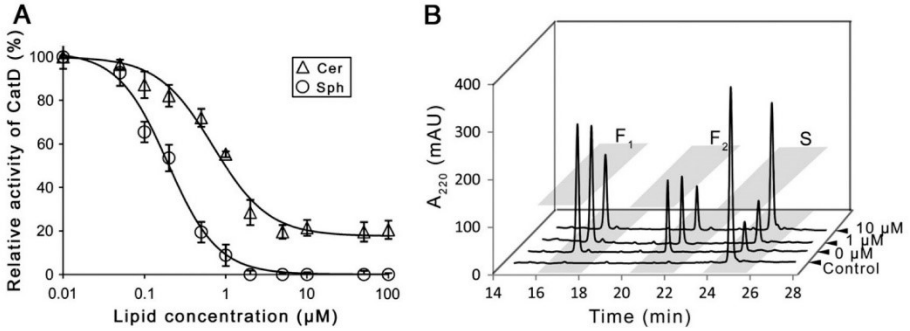


Fig. 3. Inhibition of the peptidolytic activity of CatD by sphingolipids. (A) Inhibition of CatD with varied concentrations of sphingosine (Sph; IC_{50} = 0.19 μM) and C6-ceramide **12** (Cer; IC_{50} = 1.0 μM). CatD activity was measured with a FRET-based octapeptide substrate at pH 5.5. Mean values ± SD are expressed as the remaining activity relative to the uninhibited control (100%). (B) The cleavage efficiency of CatD was inhibited in the presence of sphingosine. Synthetic peptide [15p-36p] (derived from the CatD propeptide) was digested by CatD at pH 4.2, and the reaction mixture was separated by RP-HPLC. The RP-HPLC elution profile was monitored by measuring the absorbance at 220 nm; the positions of substrate (S: [15p-36p]) and its cleavage products (fragments F_1: [27p-36p] and F_2: [15p-26p]) are indicated. The profiles are marked with the sphingosine concentration (1 μM; 10 μM; 0 μM, digest in the absence of sphingosine; Control, uncleaved substrate).

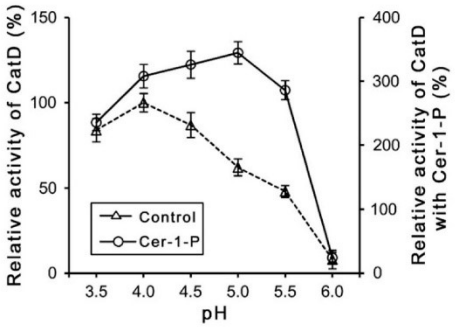


Fig. 5. pH profile of CatD determined in the absence (Control, left axis) and presence of 1 μM C8-ceramide-1-phosphate (Cer-1-P, right axis). The peptidolytic activity of CatD was measured in a FRET-based assay. Mean values ± SD are given normalized to the maximum value measured without C8-ceramide-1-phosphate (100%).

increased at higher pH values (Fig. 4). For comparison, the pH profile of CatD with a pH optimum at 4.2 is provided in Fig. 5. The reversibility of the sphingosine inhibition was examined by a dilution experiment, which showed that the activity of the mixture of CatD and sphingosine depends on the final concentration of sphingosine regardless of its preincubation concentration (see Materials and methods).

In addition to the FRET assay, we introduced an assay system based on an unmodified peptide substrate for testing the modulation of peptidolytic activity of CatD by SLs. The substrate was designed based on the structure of the CatD propeptide that is autocatalytically cleaved at the Leu26p-Ile27p bond during CatD activation [19]. The synthetic 22-mer peptide fragment [15p-36p] spanning the cleavage site was specifically hydrolyzed by CatD at the native position, and the reaction was monitored by separation of the cleavage products by HPLC and their identification by mass spectrometry (Fig. 3B). Using this assay, we demonstrated that the incubation of CatD in the presence of various concentrations of sphingosine inhibited cleavage of the peptide [15p-36p] in a SL concentration-dependent manner (Fig. 3B).

Kinetic analysis was performed to determine the mode of inhibition of CatD by SLs. The hydrolysis of the FRET peptidic substrate followed standard Michaelis-Menten kinetics with a K_m of 59 μM. The initial rates of product formation in the presence of C8-ceramine (**11**), one of the most inhibitory SLs tested, are shown in a Lineweaver-Burk plot (Fig. 6A). The family of straight lines intersects on the 1/v axis, indicating competitive inhibition ($K_i = 0.07$ μM).

Furthermore, a fluorescence polarization experiment was used to analyze the interaction of SLs with CatD; this approach allows for measurement of the inhibitor affinity independent of enzyme activity. As a fluorescent reporter probe, we employed C5-ceramide labeled with a BODIPY moiety (**20**), which inhibits CatD with an IC_{50} of 11.7 μM (Table 1). The probe was used as a primary CatD ligand in a binding experiment monitored by fluorescence anisotropy. The measurement was performed as a competition assay between the probe and sphingosine to yield an equilibrium dissociation constant (K_d) of 2.0 μM for the sphingosine-CatD complex (Fig. 6B). This experiment directly demonstrated the binding of SLs to CatD and indicated that sphingosine and ceramide-type inhibitors share the same binding site on the CatD molecule.

Finally, we investigated how the SL-mediated inhibition of CatD is influenced by a PSL activator. We compared the inhibition curves of sphingosine in the presence and absence of 10 μM C8-ceramide-1-phosphate at pH 5.0 using the FRET assay. The determined IC_{50} values were 0.32 ± 0.7 μM for sphingosine alone and 2.1 ± 0.7 μM for sphingosine combined with C8-ceramide-1-phosphate. This demonstrated that the apparent inhibitory effect was significantly reduced by the activator, indicating that SLs and PSLs are able to concurrently control the net activity of CatD.

3.3. CatD is activated by PSLs and this interaction is modulated by pH, ions, and heparin

PSLs without further substitution of the phosphate group induced the peptidolytic activity of CatD (Table 1). The activation curves of CatD measured with the FRET octapeptide substrate are shown in Fig. 7A. The induction of CatD activity increased with increasing PSL concentration of as shown for C8-ceramide-1-phosphate (**17**) and sphingosine-1-phosphate (**8**), and reached a maximum in the 100 micromolar range. The half-maximal induction was obtained with submicromolar PSL concentrations. Interestingly, the activation changed the pH profile of CatD peptidolytic activity (Fig. 4). In the presence of C8-ceramide-1-phosphate, the most activating PSL, the pH optimum of CatD was shifted about 1 pH unit higher to pH 5.0 (Fig. 5). Furthermore, an HPLC assay with the [15p-36p] peptidic substrate was employed to analyze the activation in an alternative assay system. It clearly demonstrated that incubation in the presence of C8-ceramide-1-phosphate augmented the cleavage efficiency of CatD in a PSL concentration-dependent manner (Fig. 7B).

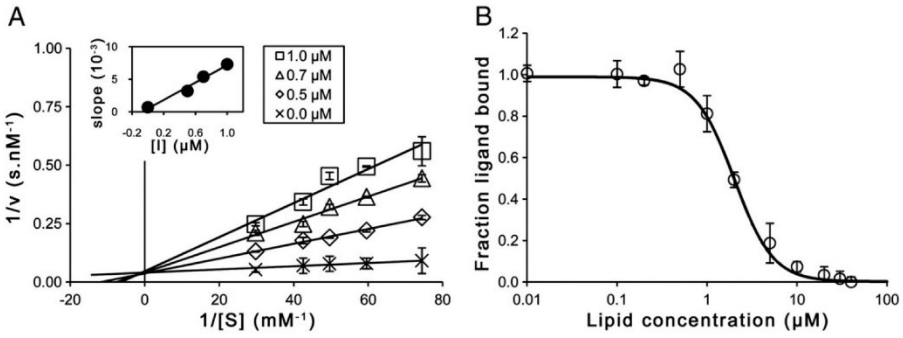


Fig. 6. Mechanism of the interaction of inhibitory sphingolipids with CatD. (A) The mode of inhibition of the peptidolytic activity of CatD by C8-ceramine was determined to be competitive using a FRET-based assay at pH 4.2. CatD activity was measured with varied concentrations of substrate (S) and inhibitor (I, concentrations are indicated). Lineweaver-Burk plot is presented together with a secondary plot of the same data (insert); $K_i = 0.07$ μM. Mean values ± SD are given. (B) Fluorescence polarization assay of the binding of sphingosine to CatD at pH 5.5 plotted as a fraction of the ligand bound. Competitive displacement of the fluorescent probe BODIPY FL C5-ceramide (2 μM) complexed with CatD (50 nM) by sphingosine (concentration indicated) revealed the optimized parameter $K_d = 2.0$ μM for the sphingosine-CatD complex.

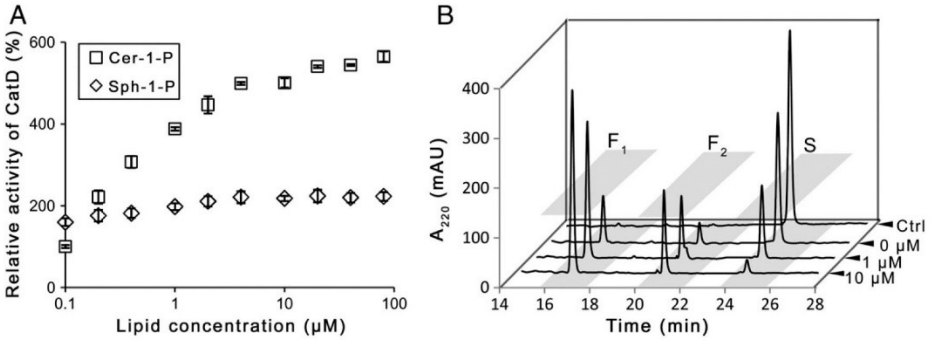


Fig. 7. Phosphosphingolipids induce the peptidolytic activity of CatD. (A) Activation of CatD with varied concentrations of sphingosine-1-phosphate (Sph-1-P) and C8-ceramide-1-phosphate (Cer-1-P). CatD activity was measured in a fluorescence assay with a specific FRET octapeptide substrate at pH 4.2. Mean values ± SD are expressed relative to the non-activated control (100%). (B) The cleavage efficiency of CatD was induced in the presence of C8-ceramide-1-phosphate. Synthetic peptide [15p-36p] (derived from the CatD propeptide) was digested by CatD at pH 4.2, and the reaction mixture was separated by RP-HPLC. The RP-HPLC elution profile was monitored by measuring absorbance at 220 nm; the positions of substrate (S: [15p-36p]) and its cleavage products (fragments F_1: [27p-36p] and F_2: [15p-26p]) are indicated. The profiles are marked with the C8-ceramide-1-phosphate concentration (1 μM; 10 μM; 0 μM, digest in the absence of C8-ceramide-1-phosphate; Control, uncleaved substrate).

To describe the activation mechanism in enzymological terms, we determined the kinetic parameters of CatD with the FRET substrate for the non-induced system and system induced with 10 μM C8-ceramide-1-phosphate. We obtained the following values: $K_m = 59 \pm 9$ μM, $k_{cat} = 54 \pm 6\ s^{-1}$, $k_{cat}/K_m = 910\ mM^{-1}.s^{-1}$ (non-induced), and $K_m = 4 \pm 0.7$ μM, $k_{cat} = 108 \pm 10\ s^{-1}$, $k_{cat}/K_m = 26967\ mM^{-1}.s^{-1}$ (induced) at pH 4.2. These data indicated that PSL increased the catalytic efficiency (k_{cat}/K_m) of CatD about 30-fold due to one order of magnitude better substrate-binding affinity (K_m) and, to a lesser extent, a faster turnover rate (k_{cat}).

Finally, we investigated the interaction between CatD and PSLs in the presence of ions and heparin. In Fig. 8A, CatD was activated by C8-ceramide-1-phosphate, and the effect of various concentrations of NaCl, $CaCl_2$, and $MgCl_2$ was compared. The PSL-mediated activation of CatD was suppressed by Ca^{2+} and Mg^{2+} by about 60% of the induced activity (Fig. 8A), although the tested salt concentrations did not significantly influence CatD activity in the absence of PSL (data not shown). The suppression was concentration-dependent, and Mg^{2+} exerted its effect at lower concentrations than Ca^{2+}. Furthermore, the combined action of C8-ceramide-1-phosphate and heparin on CatD was analyzed (Fig. 8B). Heparin is a natural sulfated glycosaminoglycan that activates CatD [25]. Here, it was applied at a saturating concentration that produced ~3-fold activation of CatD. The experiment showed that the induction effects mediated by both activators are additive at lower PSL concentrations and non-additive at saturating PSL concentrations.

4. Discussion

In this work, we discovered that SLs are complex modulators of the peptidolytic activity of CatD. SLs are capable of inhibiting or activating CatD, depending on their structure. First, a panel of SL derivatives was analyzed to identify the structural variables that determine their modulatory effects on CatD activity. Second, we investigated the mechanism of the interaction of SLs with CatD and factors that control this process. The effect of SLs was independently demonstrated in two different CatD activity assay systems using specific peptidic substrates applied in FRET-based and HPLC-based assays.

Sphingosine- and ceramide-type SLs exert inhibitory effects on CatD activity. The inhibitory activity of the sphingosine backbone was found to be relatively tolerant to structural modification. Of particular interest is the apparent unimportance of the allylic hydroxyl moiety; similar inhibitory activity was observed in sphingoid bases with this *trans*-double bond either saturated or substituted by a triple

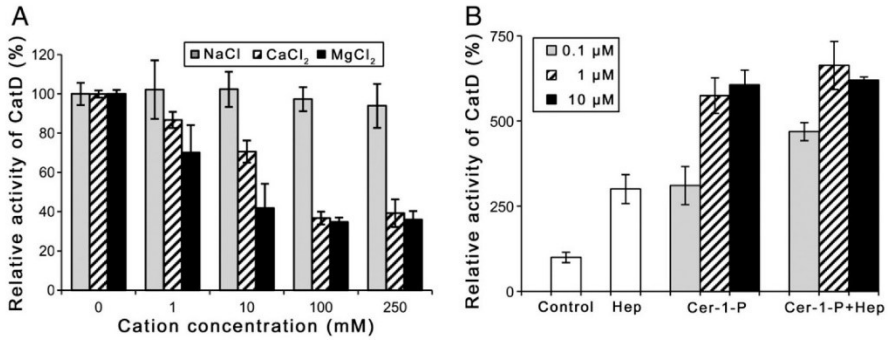


Fig. 8. Induction of CatD by C8-ceramide-1-phosphate in the presence of ions and heparin. The peptidolytic activity of CatD was measured in a FRET-based assay at pH 5.0; mean values ± SD are given. (A) CatD activity was induced by 1 μM C8-ceramide-1-phosphate in the presence of varying concentrations of NaCl, $CaCl_2$, and $MgCl_2$. The activity values are expressed relative to the controls induced in the absence of salts (100%). (B) CatD activity was induced by 0.1–10 μM C8-ceramide-1-phosphate (Cer-1-P) or 750 μg/ml heparin (Hep) or by the combined action of both compounds (Cer-1-P + Hep). The activity values are expressed relative to the non-induced control (100%).

bond. Interestingly, similar structural tolerance has been observed in the anticancer effects of the sphingoid bases [26]. On the other hand, ceramides, *i.e. N*-acylsphingosines, were more sensitive to structural (including sterochemical) alterations than the sphingoid bases. In particular, the allylic hydroxyl was found essential for the inhibitory effect of ceramides on CatD activity. This structural requirement of ceramides has also been reported for their ability to induce apoptosis [27]. Furthermore, the results revealed that the base chain length and the acyl chain length are important factors determining the ability of sphingosines and ceramides to inhibit CatD. Our data indicated that certain optimized chain lengths are necessary for effective binding to CatD, suggesting the importance of hydrophobic interactions in this process. Finally, we hypothesize that the (partial) positive charge on the nitrogen linked to the sphingoid C2 atom is important for the interaction of sphingolipids with CatD; this is suggested by the inactivity of an azido analog (**4**) compared to sphingosines and ceramides containing (alkyl)amino and amido moieties.

To study the mechanism of the inhibitory interaction of SLs with CatD, several approaches were applied. Fluorescence polarization measurements demonstrated that the SL inhibitors directly bind to the CatD molecule and that there is a common binding site for sphingosines and ceramides. Dilution experiments showed the reversibility of SL binding, and kinetic analysis indicated a competitive mode of inhibition of CatD by SLs. Using SDS PAGE analysis, we also found that the effect of SL-mediated inhibition is not associated with an autoproteolytic degradation of CatD under assay conditions (data not shown). Taken together, our results suggested that the SL inhibitors bind at the active site of CatD and compete with a peptidic substrate for its binding site. However, based on the obtained data and the inspection of the CatD crystallographic model (PDB code 1LYB), the exact position of this binding site cannot be unambiguously predicted. The active site of CatD contains subsites that accommodate hydrophobic parts of substrates, which is reflected in the enzyme cleavage specificity [28]. This was also demonstrated by the peptidic substrates of CatD used in this study, which were cleaved between the hydrophobic amino acids at the bonds Phe-Nph and Leu-Ile (in the FRET and [15p-36p] substrates, respectively). Therefore, we may assume that SLs as amphiphilic ligands can occupy, at least in part, such binding subsites. Nevertheless, to fully understand the binding mode of SLs to CatD, the crystal structure of a SL-CatD complex will be required. To that end, we are pursuing the co-crystallization of CatD with the most potent SL inhibitors.

In contrast to the SLs, specific PSLs were found to activate CatD. In particular, phosphorylation of the primary hydroxyl of both sphingosine and ceramide resulted in induction of the peptidolytic activity of CatD. However, the activation effect was also found for phosphatidic acid, which suggested that the induction of CatD activity could be attributed more generally to phospholipids with an unsubstituted monoester phosphate moiety. This is in line with an earlier finding that acidic lipids can enhance CatD proteolysis [29]. Using enzyme kinetics, we analyzed the activation effect for C8-ceramide-1-phosphate, the most activating compound, and found that the activation is based predominantly on the improved substrate-binding affinity. Furthermore, the activation mediated by C8-ceramide-1-phosphate was about two-fold higher than that mediated by heparin, a sulfated polysaccharide that has been reported to activate CatD [25]. The analysis of the combined action of C8-ceramide-1-phosphate and heparin suggested that these negatively-charged activators might interact with different binding sites on the CatD molecule. The exact binding site for heparin is unknown, but this interaction most likely occurs through positively-charged surface region(s) of CatD, as similarly reported for cysteine cathepsins [30,31]. Indeed, the CatD crystallographic model shows that there are four basic patches in the vicinity of the active site, and we hypothesize that some of these candidate sites are utilized for interaction with PSLs. This interaction can be potentially multivalent due to the macromolecular character of heparin and the micelle-forming ability of PSLs. The importance of the negatively-charged phosphate group of PSLs for the interaction with CatD was shown by the specific effect of Mg^{2+}/Ca^{2+}, which substantially suppressed the induction effect. This can be explained by the effective coordination of the divalent cations to the phosphate group, thus masking the phosphate functionality of PSLs and preventing their binding to CatD. Our results do not exclude the possibility that the activating PSLs might interact with CatD through a different mode of recognition than the inhibitory SLs, although their binding sites are not functionally independent, as suggested by the experiment combining SP and PSL action.

To evaluate the SL/PSL-mediated modulation of CatD in a physiological context, future work will be needed. We suppose that the modulation is physiologically relevant since CatD and SLs/PSLs can co-localize based on their broad intracellular and extracellular distribution. Furthermore, SLs/PSLs were effective *in vitro* at physiological concentrations (generally up to the low micromolar level). In addition to an inhibition/activation, we also demonstrated that the presence of PSLs causes a shift in the CatD pH optimum to a higher pH. This shift may be important for CatD activity when the enzyme is translocated outside lysosomes, *e.g.* during pathophysiologic conditions [8,9,14]. SLs/PSLs can be also involved in the modulation of autocatalytic activation of the CatD zymogen, as suggested by their effect on the cleavage of the synthetic processing region from the CatD propeptide. The polarity of the CatD response to SLs as inhibitors and the role of PSLs as activators deserve special attention. Bioactive SLs such as ceramide and sphingosine have been implicated in promoting apoptosis and cell cycle arrest, while PSLs (ceramide-1-phosphate and sphingosine-1-phosphate) are antagonistic molecules with generally antiapoptotic and mitogenic properties [3,4,12,32]. CatD is involved in both apoptosis and mitogenesis [8,9,12]. Thus, it is intriguing to speculate that the role of CatD in these processes is under complex control of the regulatory SL-PSL system and that CatD is evolutionarily designed to sense these modulators. In conclusion, we discovered that CatD is effectively modulated by naturally occurring SLs/PSLs as well as their specific synthetic derivatives. Based on this, we propose that SLs/PSLs might serve as a novel scaffold for the development of selective regulators of CatD-dependent processes with potential therapeutic applications.

Supplementary materials related to this article can be found online at doi:10.1016/j.bbalip.2011.09.005.

Acknowledgements

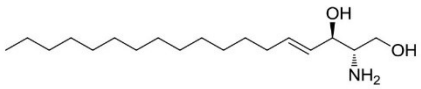
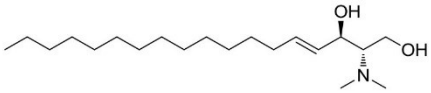
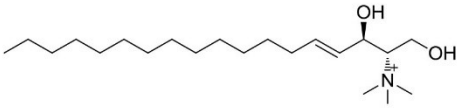
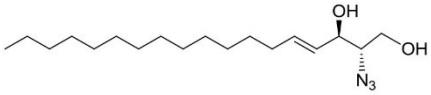
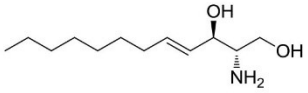
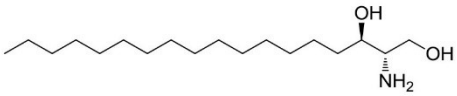
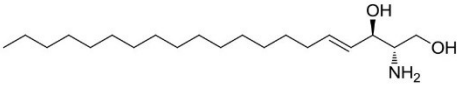
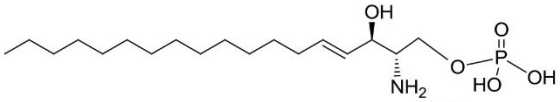
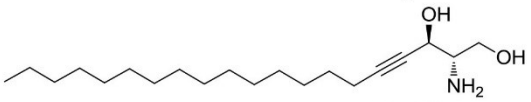
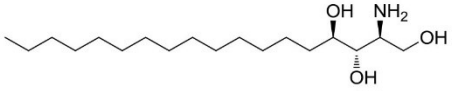
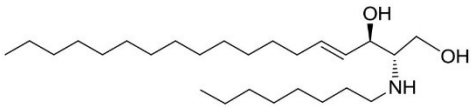
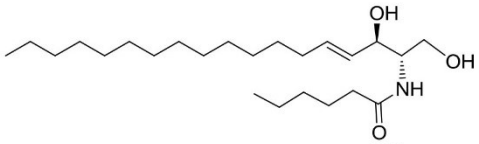
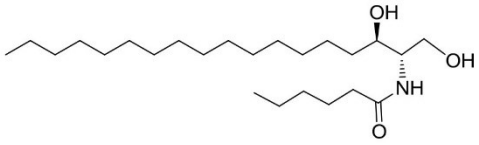
This work was supported by grant no. IAA400550705 from the Grant Agency of the Academy of Sciences of the Czech Republic and by research project Z40550506. The authors thank Hillary Hoffman for critical reading of the manuscript.

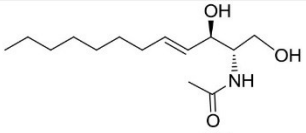
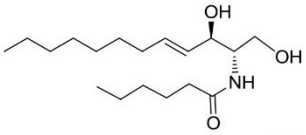
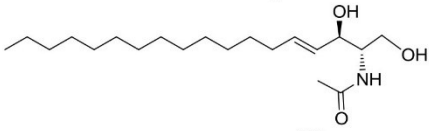
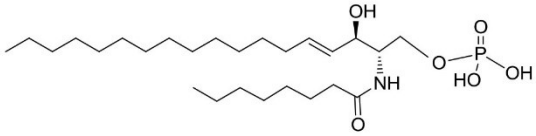
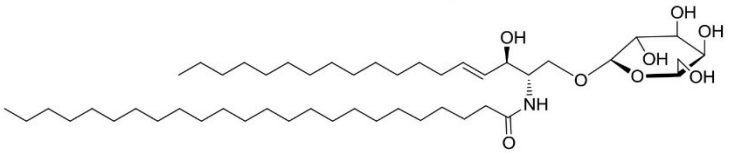
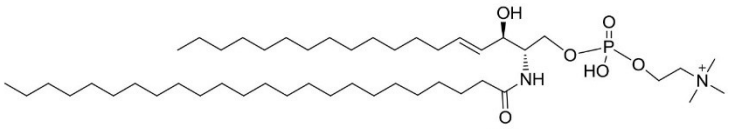
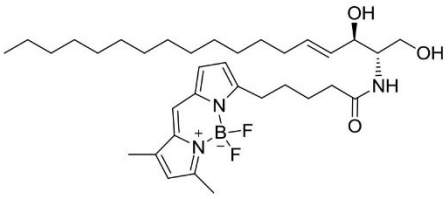
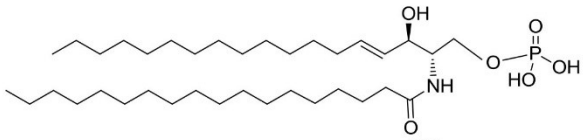
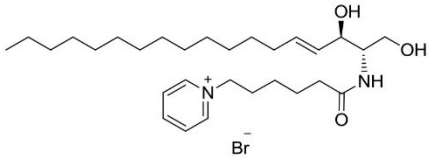
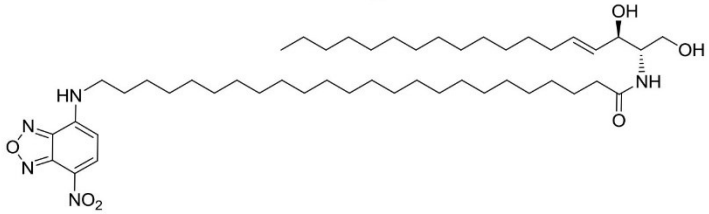
References

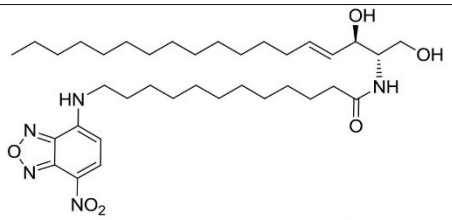
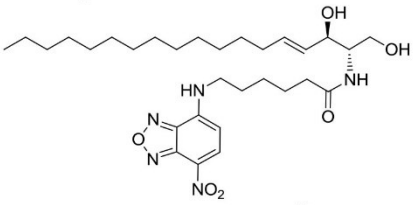
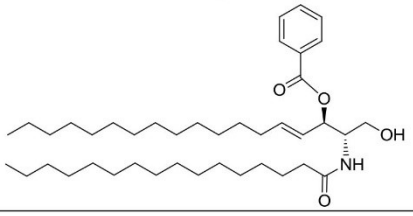
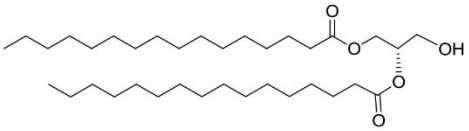
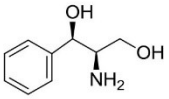
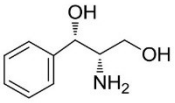
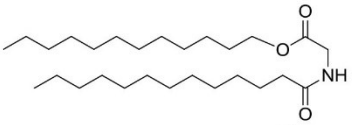
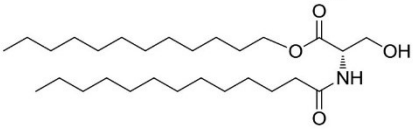
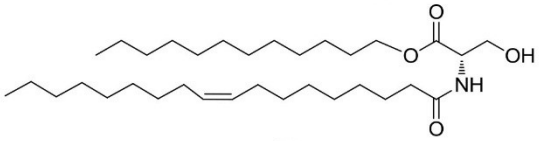
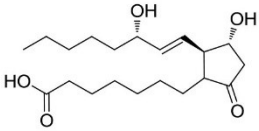
[1] N. Bartke, Y.A. Hannun, Bioactive sphingolipids: metabolism and function, J. Lipid Res. 50 (Suppl) (2009) S91–S96.
[2] Y.A. Hannun, L.M. Obeid, Principles of bioactive lipid signalling: lessons from sphingolipids, Nat. Rev. Mol. Cell Biol. 9 (2008) 139–150.
[3] B. Ogretmen, Sphingolipids in cancer: regulation of pathogenesis and therapy, FEBS Lett. 580 (2006) 5467–5476.
[4] A. Morales, J.C. Fernandez-Checa, Pharmacological modulation of sphingolipids and role in disease and cancer cell biology, Mini Rev. Med. Chem. 7 (2007) 371–382.
[5] C.F. Snook, J.A. Jones, Y.A. Hannun, Sphingolipid-binding proteins, Biochim. Biophys. Acta 1761 (2006) 927–946.
[6] M. Heinrich, M. Wickel, W. Schneider-Brachert, C. Sandberg, J. Gahr, R. Schwandner, T. Weber, P. Saftig, C. Peters, J. Brunner, M. Kronke, S. Schutze, Cathepsin D targeted by acid sphingomyelinase-derived ceramide, EMBO J. 18 (1999) 5252–5263.
[7] P. Saftig, M. Hetman, W. Schmahl, K. Weber, L. Heine, H. Mossmann, A. Koster, B. Hess, M. Evers, K. von Figura, C. Peters, Mice deficient for the lysosomal proteinase cathepsin D exhibit progressive atrophy of the intestinal mucosa and profound destruction of lymphoid cells, EMBO J. 14 (1995) 3599–3608.
[8] O. Masson, A.S. Bach, D. Derocq, C. Prebois, V. Laurent-Matha, S. Pattingre, E. Liaudet-Coopman, Pathophysiological functions of cathepsin D: targeting its catalytic activity versus its protein binding activity? Biochimie 92 (2010) 1635–1643.

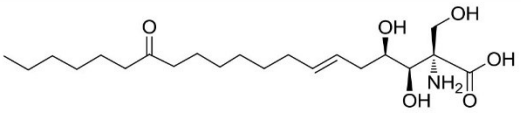
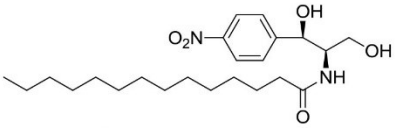
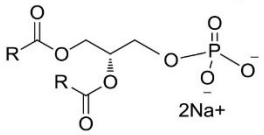
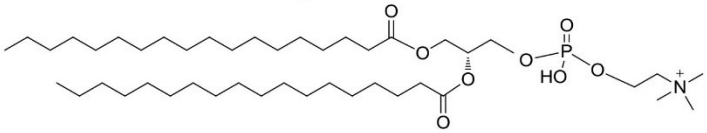
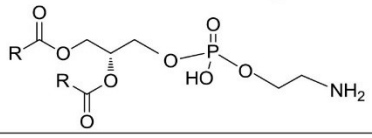
[9] P. Benes, V. Vetvicka, M. Fusek, Cathepsin D—many functions of one aspartic protease, Crit. Rev. Oncol. Hematol. 68 (2008) 12–28.
[10] J. Rodriguez, J. Vazquez, M.D. Corte, M. Lamelas, M. Bongera, M.G. Corte, A. Alvarez, M. Allende, L. Gonzalez, M. Sanchez, M. Vijande, M.J. Garcia, F. Vizoso, Clinical significance of cathepsin D concentration in tumor cytosol of primary breast cancer, Int. J. Biol. Markers 20 (2005) 103–111.
[11] N. Bossard, F. Descotes, A.G. Bremond, Y. Bobin, H.P. De Saint, F. Golfier, A. Awada, P.M. Mathevet, L. Berrerd, Y. Barbier, J. Esteve, Keeping data continuous when analyzing the prognostic impact of a tumor marker: an example with cathepsin D in breast cancer, Breast Cancer Res. Treat. 82 (2003) 47–59.
[12] E. Liaudet-Coopman, M. Beaujouin, D. Derocq, M. Garcia, M. Glondu-Lassis, V. Laurent-Matha, C. Prebois, H. Rochefort, F. Vignon, Cathepsin D: newly discovered functions of a long-standing aspartic protease in cancer and apoptosis, Cancer Lett. 237 (2006) 167–179.
[13] M.M. Gopalakrishnan, H.W. Grosch, S. Locatelli-Hoops, N. Werth, E. Smolenova, M. Nettersheim, K. Sandhoff, A. Hasilik, Purified recombinant human prosaposin forms oligomers that bind procathepsin D and affect its autoactivation, Biochem. J. 383 (2004) 507–515.
[14] M. Beaujouin, C. Prebois, D. Derocq, V. Laurent-Matha, O. Masson, S. Pattingre, P. Coopman, N. Bettache, J. Grossfield, R.E. Hollingsworth, H. Zhang, Z. Yao, B.T. Hyman, P. van der Geer, G.K. Smith, E. Liaudet-Coopman, Pro-cathepsin D interacts with the extracellular domain of the beta chain of LRP1 and promotes LRP1-dependent fibroblast outgrowth, J. Cell Sci. 123 (2010) 3336–3346.
[15] Y. Zhang, X. Li, K.A. Becker, E. Gulbins, Ceramide-enriched membrane domains—structure and function, Biochim. Biophys. Acta 1788 (2009) 178–183.
[16] M. Heinrich, J. Neumeyer, M. Jakob, C. Hallas, V. Tchikov, S. Winoto-Morbach, M. Wickel, W. Schneider-Brachert, A. Trauzold, A. Hethke, S. Schutze, Cathepsin D links TNF-induced acid sphingomyelinase to Bid-mediated caspase-9 and −3 activation, Cell Death Differ. 11 (2004) 550–563.
[17] D. De Stefanis, P. Reffo, G. Bonelli, F.M. Baccino, G. Sala, R. Ghidoni, P. Codogno, C. Isidoro, Increase in ceramide level alters the lysosomal targeting of cathepsin D prior to onset of apoptosis in HT-29 colon cancer cells, Biol. Chem. 383 (2002) 989.
[18] M. Fusek, M. Baudys, P. Metcalf, Purification and crystallization of human cathepsin D, J. Mol. Biol. 226 (1992) 555–557.
[19] M. Masa, L. Maresova, J. Vondrasek, M. Horn, J. Jezek, M. Mares, Cathepsin D propeptide: mechanism and regulation of its interaction with the catalytic core, Biochemistry 45 (2006) 15474–15482.
[20] J. Novotny, K. Pospechova, A. Hrabalek, R. Cap, K. Vavrova, Synthesis of fluorescent C24-ceramide: evidence for acyl chain length dependent differences in penetration of exogenous NBD-ceramides into human skin, Bioorg. Med. Chem. Lett. 19 (2009) 6975–6977.
[21] J. Novotny, B. Janusova, M. Novotny, A. Hrabalek, K. Vavrova, Short-chain ceramides decrease skin barrier properties, Skin Pharmacol. Physiol. 22 (2009) 22–30.
[22] K. Vavrova, A. Hrabalek, P. Dolezal, L. Samalova, K. Palat, J. Zbytovska, T. Holas, J. Klimentova, Synthetic ceramide analogues as skin permeation enhancers: structure-activity relationships, Bioorg. Med. Chem. 11 (2003) 5381–5390.
[23] J.R. Lakowics, Principles of Fluorescence Spectroscopy, second ed., Springer, New York, 1999.
[24] J. Ouveley, O.W. Thiele, A note on the spectrophotometric determination of long chain bases in lipids, Lipids 15 (1979) 194–195.
[25] M. Beckman, C. Freeman, C.R. Parish, D.H. Small, Activation of cathepsin D by glycosaminoglycans, FEBS J. 276 (2009) 7343–7352.
[26] J.M. Padron, Sphingolipids in anticancer therapy, Curr. Med. Chem. 13 (2006) 755.
[27] N.S. Radin, Designing anticancer drugs via the achilles heel: ceramide, allylic ketones, and mitochondria, Bioorg. Med. Chem. 11 (2003) 2123–2142.
[28] B.M. Dunn, S. Hung, The two sides of enzyme-substrate specificity: lessons from the aspartic proteinases, Biochim. Biophys. Acta 1477 (2000) 231–240.
[29] K.R. Williams, N.D. Williams, W. Konigsberg, R.K. Yu, Acidic lipids enhance cathepsin D cleavage of the myelin basic protein, J. Neurosci. Res. 15 (1986) 137–145.
[30] M. Fairhead, S.M. Kelly, C.F. van der Walle, A heparin binding motif on the prodomain of human procathepsin L mediates zymogen destabilization and activation, Biochem. Biophys. Res. Commun. 366 (2008) 862–867.
[31] D. Caglic, J.R. Pungercar, G. Pejler, V. Turk, B. Turk, Glycosaminoglycans facilitate procathepsin B activation through disruption of propeptide-mature enzyme interactions, J. Biol. Chem. 282 (2007) 33076–33085.
[32] A. Gomez-Munoz, Ceramide 1-phosphate/ceramide, a switch between life and death, Biochim. Biophys. Acta 1758 (2006) 2049–2056.

Supplemental Table S1
Structures of the sphingolipids and related compounds presented in Table 1.

Comp. no.	Abbreviation	Structure
	Sphingosine derivatives	
1	sphingosine	
2	*N,N*-dimethylsphingosine	
3	*N,N,N*-trimethylsphingosine	
4	azidosphingosine	
5	C12-sphingosine	
6	dihydrosphingosine (sphinganine)	
7	C20-sphingosine	
8	sphingosine-1-phosphate	
9	(*2S,3R*)-2-aminoicos-4-yne-1,3-diol (triple bond C20-sphingosine)	
10	phytosphingosine	
11	C8-ceramine (*N*-octylsphingosine)	
	Ceramide derivatives	
12	C6-ceramide (C18 base)	
13	C6-dihydroceramide (C18 base)	

14	C2-ceramide (C12 base)	
15	C6-ceramide (C12 base)	
16	C2-ceramide (C18 base)	
17	C8-ceramide-1-phosphate	
18	1-*O*- β-D-galactosyl-ceramide[a]	
19	sphingomyelin	
20	BODIPY FL C5-ceramide	
21	C18-ceramide-1-phosphate[a]	
22	C6-pyridinium-ceramide	
23	C24-NBD-ceramide	

24	C12-NBD-ceramide	
25	C6-NBD-ceramide	
26	3-*O*-benzoyl-C16-ceramide	
	Others	
27	1,2-dipalmitoyl-glycerol	
28	(*1R,2R*)-2-amino-1-phenyl-1,3-propanediol	
29	(*1S,2S*)-2-amino-1-phenyl-1,3-propanediol	
30	*N*-dodecanoylglycine dodecyl ester	
31	*N*-dodecanoyl-L-serine dodecyl ester	
32	*N*-oleoyl-L-serine dodecyl ester	
33	prostaglandin E1	

34	myriocin	
35	(*1R,2R*)-2-myristylamino-1-(4-nitrophenyl)-1,3-propandiol	
36	phosphatidic acid[b]	
37	phosphatidylcholine	
38	phosphatidylethanolamine[b]	

[a]Major acyl component of natural mixtures from bovine brain is presented.

[b]Natural mixtures from egg yolk.

4.2 Publikace č. 2: Novel structural mechanism of allosteric regulation of aspartic peptidases via an evolutionarily conserved exosite

4.2.1 Souhrn

Proenzymy (zymogeny) aspartátových proteas mají na N-konci molekuly tzv. propeptid, který je během procesu aktivace odštěpen za vzniku zralého, katalyticky aktivního enzymu. Tato studie byla zaměřená na strukturně funkční analýzu propeptidu u proIrKatD1, katepsinu D1 z klíštěte obecného (*Ixodes ricinus*). Propeptid IrKatD1 se liší od propeptidů ostatních aspartátových proteas A1 rodiny pepsinu svojí malou velikostí (má pouze 22 aminokyselin oproti téměř dvojnásobku u lidského KatD). Pomocí dále uvedených přístupů bylo zjištěno, že zjednodušená stavba propeptidu proIrKatD1 je plně funkční a představuje tak ideální prototyp pro studium vlastností této domény u aspartátových proteas obecně.

Detailně byl popsán proces autoaktivace proenzymu IrKatD1 (proIrKatD1) na zralý aktivní enzym, ke kterému dochází v prostředí kyselého pH. Jedná se o monomolekulární reakci, kdy je nejprve odštěpen celý propeptid, který je následně dále fragmentován na dvě části. Výchozí a konečný stav konverze byl popsán pomocí krystalových struktur jak proenzymu, tak zralého enzymu. Prostorová struktura proIrKatD1 umožnila podrobnou analýzu interakce intaktního propeptidu s katalytickou doménou. Propeptid blokuje přístup substrátu do aktivního centra, zároveň je ale dostatečné vzdálen od katalytických aspartátů, aby nedošlo k jeho nekontrolovanému štěpení. N-koncový segment propeptidu má konformaci ß-řetězce, který tvoří několik vodíkových můstků s jádrem enzymu a váže se na povrch enzymu v místě, které bylo popsáno jako exomísto vzhledem k jeho umístění mimo vlastní aktivní centrum. Jednotlivé vazebné pozice exomísta (označené jako S_1^E až S_8^E) umožňují interakci s osmi zbytky N-konce propeptidu.

Jedním z nejdůležitějších výsledků této práce bylo zjištění, že N-koncový oligopeptidový fragment, který je generovaný z propeptidu během autoaktivace, je schopen autonomně fungovat jako silný nanomolární inhibitor (IC_{50} = 12 nM) rodičovského enzymu. Inhibice byla popsaná pomocí kinetických měření jako smíšená nekompetitivní inhibice, která je závislá na pH, a pomocí termoforetických měření byla zjištěna disociační konstanta. Dála byla určena krystalová struktura inhibičního fragmentu v komplexu s IrKatD1, jenž prokázala, že se inhibitor váže do exomísta a indukuje alosterickou

konformační změnu v aktivním centru IrKatD1, které se tak stane nedostupným pro substrát. Mapování inhibičního fragmentu pomocí sady delečních syntetických peptidů umožnilo identifikovat kritický inhibiční motiv tvořený aminokyselinami 1-8 (číslování podle proenzymu). Příspěvek jednotlivých zbytků k interakci byl studován pomocí alaninových substitucí, které ukazují zásadní roli leucinu v pozici 6; tento zbytek je také vysoce konzervovaný v sekvencích propeptidů homologických proteas.

Práce přináší objev prvního peptidového endogenního inhibitoru aspartátových proteas u vyšších organismů a podobně charakterizuje jeho mechanismus interakce. Vzhledem k sekvenční a strukturní homologii propeptidů a exomíst je pravděpodobné, že vznik a působení tohoto typu inhibitorů je u pepsinových proteas evolučně konzervované. Dále studie definuje exomísto jako molekulární cíl pro vývoj nových alosterických inhibitorů medicinálně významných aspartátových proteas.

Můj podíl na práci zahrnoval: (1) přípravu vybraných expresních vektorů mutantní formy proIrKatD1 pro expresi v bakteriálních buňkách, (2) expresi a purifikaci proIrKatD1 a IrKatD1, (3) biochemickou analýzu aktivace proIrKatD1 na zralý enzym, (4) určení inhibičních konstant IC_{50} pro syntetické fragmenty odvozené z propeptidu proIrKatD1, (5) analýzu stability připravených inhibitorů v přítomnosti zralého IrKatD1 pomocí HPLC, (6) přípravu krystalů zralého IrKatD1, jeho inhibičních komplexů a proIrKatD1 pro rentgenostrukturní analýzu, (7) analýzu intermolekulárních interakcí v prostorových modelech a molekulární grafiku, (8) grafické zpracování dat a přípravu manuskriptu.

CellPress

Cell Chemical Biology
Article

Novel Structural Mechanism of Allosteric Regulation of Aspartic Peptidases via an Evolutionarily Conserved Exosite

Iva Hánová,[1,2] Jiří Brynda,[1] Radka Houštecká,[1,3] Nawsad Alam,[5] Daniel Sojka,[4] Petr Kopáček,[4] Lucie Marešová,[1] Jiří Vondrášek,[1] Martin Horn,[1] Ora Schueler-Furman,[5] and Michael Mareš[1,6,*]

[1]Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Prague, Czech Republic
[2]Department of Biochemistry, Faculty of Science, Charles University, 12840 Prague, Czech Republic
[3]First Faculty of Medicine, Charles University, 12108 Prague, Czech Republic
[4]Institute of Parasitology, Biology Centre of the Czech Academy of Sciences, 37005 Ceske Budejovice, Czech Republic
[5]Department of Microbiology and Molecular Genetics, Institute for Biomedical Research IMRIC, Hebrew University, Hadassah Medical School, 91120 Jerusalem, Israel
[6]Lead Contact
*Correspondence: mares@uochb.cas.cz
https://doi.org/10.1016/j.chembiol.2018.01.001

SUMMARY

Pepsin-family aspartic peptidases are biosynthesized as inactive zymogens in which the propeptide blocks the active site until its proteolytic removal upon enzyme activation. Here, we describe a novel dual regulatory function for the propeptide using a set of crystal structures of the parasite cathepsin D IrCD1. In the IrCD1 zymogen, intramolecular autoinhibition by the intact propeptide is mediated by an evolutionarily conserved exosite on the enzyme core. After activation, the mature enzyme employs the same exosite to rebind a small fragment derived from the cleaved propeptide. This fragment functions as an effective natural inhibitor of mature IrCD1 that operates in a pH-dependent manner through a unique allosteric inhibition mechanism. The study uncovers the propeptide-binding exosite as a target for the regulation of pepsin-family aspartic peptidases and defines the structural requirements for exosite inhibition.

INTRODUCTION

Aspartic peptidases of the pepsin family (A1 family in the MEROPS database) are broadly distributed in eukaryotic organisms. Human family members are the most well studied, including the gastric enzymes pepsin and chymosin; the blood pressure regulator renin; cathepsins D and E, involved in several pathologies, including cancer; and β-secretase (BACE), implicated in Alzheimer disease (Rawlings and Barrett, 2013). Despite the importance of pepsin-family peptidases, little is known about the molecules that function as physiological regulators of their activity. There are no known specific endogenous inhibitors of human pepsin-family peptidases, other than recently discovered sphingolipid inhibitors of human cathepsin D (Zebrakovska et al., 2011). Only a few proteinaceous compounds that might serve as endogenous inhibitors in lower organisms have been identified (Lenarcic and Turk, 1999; Li et al., 2000).

Therefore, it is generally accepted that the major mechanism regulating the activity of pepsin-family peptidases is activation of the proenzyme. These peptidases are biosynthesized in the form of inactive proenzymes (zymogens), in which the active site is blocked by a propeptide domain located at the N terminus of the molecule. The propeptide is cleaved off at acidic pH either by autocatalytic processing (via a monomolecular or bimolecular reaction) or processing assisted by another peptidase, leading to active mature enzyme (for review, see Dunn, 2002; Khan et al., 1999; Richter et al., 1998). So far, several crystal structures of zymogens and their activation intermediates have helped increase understanding of propeptide interactions and the process of zymogen activation (James and Sielecki, 1986; Kervinen et al., 1999; Khan et al., 1997; Lee et al., 1998; Moore et al., 1995; Morales et al., 2012; Ostermann et al., 2004).

In this work, we investigated the cathepsin D-like aspartic peptidase IrCD1 from the hard tick *Ixodes ricinus*, a blood-feeding parasite that can transmit Lyme disease and tick-borne encephalitis. IrCD1 plays a critical role in triggering the proteolytic degradation pathway responsible for the digestion of host hemoglobin in tick gut cells (Horn et al., 2009). Previously, the IrCD1 propeptide was reported to be much shorter than propeptides of other pepsin-family peptidases (Sojka et al., 2012). This prompted us to investigate its functional competence in inactivating the IrCD1 zymogen. Using structural analysis of the zymogen, we found that the IrCD1 propeptide retains full regulatory capability, which is associated with a small evolutionarily conserved segment. Importantly, we identified an additional function for this propeptide segment. After it is released by fragmentation of the propeptide during autoactivation, it is able to rebind and effectively inhibit mature IrCD1.

Previously, several reports showed that propeptide-derived products of the proteolytic activation event in pepsin-family peptidases possess inhibitory activity (for references, see Masa et al., 2006); however, the molecular mechanism of this phenomenon remained unknown. Here, we demonstrate for

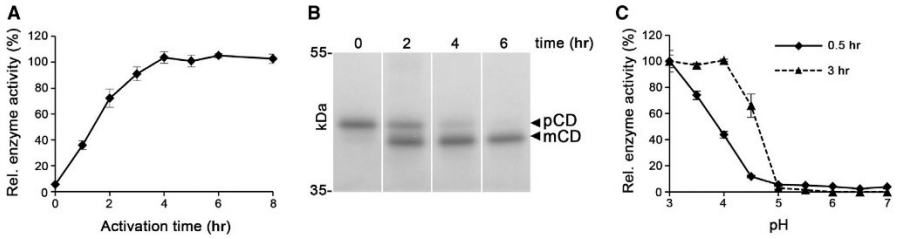


Figure 1. Autoactivation of the IrCD1 Zymogen

(A) Time course of the autoactivation. The IrCD1 zymogen was incubated at pH 4.0, and the reaction mixture was analyzed at different time points. Proteolytic activity generated during the autoactivation was determined in a kinetic assay with an FRET-peptide substrate. Mean values ±SE of triplicates are expressed relative to the maximum value. Rel., relative.

(B) The autoprocessing mixture analyzed in (A) was resolved by SDS-PAGE and visualized by protein staining. The positions for the IrCD1 zymogen (pCD) and mature enzyme (mCD) are indicated.

(C) pH profile of the autoactivation. The IrCD1 zymogen was incubated (for 0.5 or 3 hr) at various pH values, and the activity was assayed at pH 4.0 as described in (A). Mean values ±SE of triplicates are expressed relative to the maximum value for the indicated incubation time.

the first time that a free propeptide fragment interacts with an exosite located outside the active-site cleft, and we show at the structural level that this fragment inhibits the parental enzyme through a novel allosteric mechanism. This work provides insights into a new class of specific endogenous regulators of pepsin-family peptidases that are generated by autocatalytic processing and have an oligopeptide structure. Furthermore, we propose the propeptide-binding exosite, which is evolutionarily conserved, as a novel target for inhibitors of various medically important pepsin-family peptidases.

RESULTS

IrCD1 Zymogen Is Autoactivated by Propeptide Cleavage

We produced recombinant IrCD1 zymogen in a bacterial expression system. Autocatalytic activation of the zymogen was investigated using two different assays: SDS-PAGE visualization of the processed mature form and kinetic measurement of its proteolytic activity. Autoprocessing of the zymogen (47 kDa) was effective at acidic pH (below pH 5.0) and generated the enzymatically active 44-kDa mature form (Figures 1A–1C). We employed several approaches to examine the molecularity of the autoprocessing reaction. The reaction rate was independent of the zymogen concentration (Figure S1A), and it was not substantially suppressed by the interaction of the generated mature enzyme with a hemoglobin substrate or the specific macromolecular inhibitor, potato cathepsin D inhibitor (PDI) (Figures S1B and S1C). These findings indicated that the critical process in formation of the mature enzyme has a unimolecular character. This conclusion was further supported by analysis of the active-site mutated (D249N) zymogen incapable of autoprocessing (Figure S1D). The D249N mutant can serve as a substrate for the mature enzyme (added in a catalytic amount); however, the low rate of this trans-processing reaction (compared with autoprocessing) suggested that the bimolecular mechanism is not an important contributor to the autoactivation pathway of the wild-type zymogen. Furthermore, we found that autoactivation of the IrCD1 zymogen can be influenced by sulfated polysaccharides, the presence of which greatly suppressed the autoactivation reaction in our experimental setup (Figures S1E–S1G). This suggests that interactions with sulfated polysaccharides, which are widespread in all animal tissues, regulate the autoactivation process *in vivo*.

During autoactivation of the IrCD1 zymogen, the 22-residue propeptide is proteolytically removed, generating the mature N-terminal Ile23 (zymogen numbering) as determined by Edman sequencing. Mass spectrometry analysis of proteolytic products of the autoactivation reaction (Figure S1H) revealed that the released full-length propeptide is subsequently cleaved at the Phe9-Lys10 bond located between secondary structure elements of the propeptide. Simulated fragmentation of the synthetic propeptide by mature IrCD1 also resulted in processing at the Phe9-Lys10 site (Figure S1H).

Structure of Mature IrCD1 and Its Active Site

We solved two crystal structures of mature IrCD1: the empty enzyme and a complex with the substrate-mimicking inhibitor pepstatin. Crystallographic details are given in the STAR Methods and Table S1.

Our structure of uncomplexed IrCD1 solved at 1.88 Å resolution contains 339 amino acid residues starting with N-terminal Ile23. IrCD1 is a single polypeptide chain that adopts the classical fold of pepsin-family peptidases in which the molecule is divided into two topologically similar β-barrel-like domains, called the N- and C-terminal domains (Figure S2A). The active-site cleft containing catalytic residues Asp58 and Asp249 is located between these domains; the bottom of the cleft is covered by a typical six-stranded β sheet. The molecule is stabilized by three conserved disulfide bridges. Comparison of the IrCD1 structure with the structures of pepsin and human cathepsin D showed a high degree of similarity (root-mean-square deviation values for the structural alignment were 1.10 Å and 0.94 Å, respectively) (Figure S2B). The major differences in backbone superposition are located at the surface loop segments of IrCD1, including residues 230–236 (with a 5-residue insertion) and 327–333 (the so-called "polyproline loop" with a 3-residue deletion). The internal processing loop

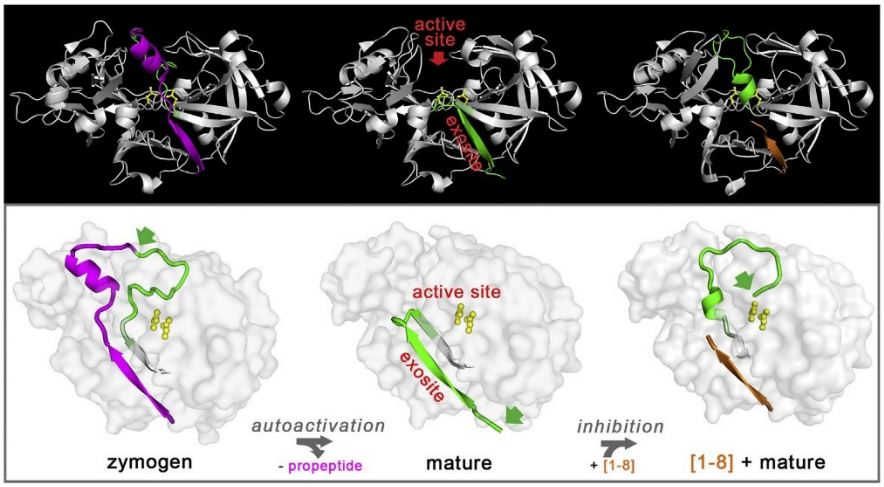


Figure 2. Proteolysis- and Ligand-Induced Conformational Changes in IrCD1

Upper : overall crystal structures of the IrCD1 zymogen (left), mature IrCD1 alone (middle), and mature IrCD1 in complex with a propeptide-derived inhibitor (right). The three-dimensional structures are shown in cartoon representation with the intact propeptide in magenta, the propeptide-derived inhibitor in orange, and the mature N terminus in green; the enzyme core is in gray. Two catalytic aspartate residues are depicted in yellow ball-and-stick representation.

Lower : side view of the structures; the enzyme core is shown as a transparent surface. Positions of the active-site cleft and the exosite are indicated (in the middle structure only, for simplicity). The autoactivation reaction scheme illustrates the conversion of the inactive zymogen into the mature enzyme by removal of the propeptide. Note that the same exosite binds the intact propeptide in the zymogen and the newly formed N terminus in the mature enzyme. The mature enzyme can be effectively inhibited by a short propeptide-derived fragment (residues 1–8) targeting the exosite. Upon binding of [1–8] to the mature enzyme, the N terminus is displaced into the active site, resulting in inhibition. The green arrows mark the changing position of the intact cleavage site and the mature N terminus generated by its proteolytic processing.

(around residue 118) characteristic of mammalian cathepsin D is absent in IrCD1.

The structure of mature IrCD1 in complex with pepstatin was solved at 1.46 Å resolution. The binding of the inhibitor into the active site induces small conformational changes in the "tyrosine flap" (a β-hairpin structure containing conserved Tyr101) and polyproline loop, which move toward the inhibitor (Figures S2). The pepstatin complex is stabilized by a set of hydrogen bonds between the backbone atoms of the inhibitor and both the main-chain and side-chain atoms of IrCD1 (Figure S2C). Two tyrosine flap residues (Gly102 and Ser103) are involved. The central statin hydroxyl group occupies the position of a water molecule that interacts with both catalytic aspartate residues in the empty enzyme structure. In general, the inhibitor side chains fill subsites S4 to S3' (except S1′, for which there is no P1′ side-chain substituent in pepstatin) and make numerous contacts with IrCD1 (listed in Table S2). The conformation of pepstatin bound to IrCD1 and its interaction network are very similar to those observed when the inhibitor is bound to human cathepsin D or pepsin (Figure S2D), indicating that the IrCD1 subsites have similar architecture for substrate binding.

Zymogen Conversion to Mature Enzyme Exploits a Propeptide-Binding Exosite

We used the structure of the inactive (D249N) IrCD1 zymogen determined at 2.30 Å resolution to describe the autoinhibition mechanism in structural terms. The zymogen contains 361 amino acid residues starting with the N-terminal Ala1 of the propeptide; the D249N mutation prevents zymogen autoprocessing. The 22-residue propeptide is situated between the two β-barrel domains of the molecule and wraps around the enzyme core (Figure 2). It consists of an N-terminal β strand followed by a single short, two-turn α helix. The C-terminal region of the propeptide is located at the top of the molecule, with the cleavage site separating it from the mature N-terminal region. Both regions adopt an Ω-shaped conformation above the active site, leading to its inactivation (Figures 2 and 3). The binding site of the propeptide N terminus and the active site are critical for intramolecular inhibition of the IrCD1 zymogen.

The N-terminal β strand of the propeptide (Phe2–Arg8) is involved in the formation of a six-stranded β sheet, making several main-chain hydrogen bond interactions with the neighboring strand, while the Arg3 side chain forms a salt bridge with Glu193 and Asp200 (Figure 4A). The hydrophobic side chains of Phe2, Ile4, and Leu6 and the hydrophobic parts of the side chains of the other residues participate in the formation of an internal hydrophobic core (Table S3, Figure S3). Importantly, the N-terminal β strand of the propeptide forms a similar hydrogen bonding network as the N-terminal segment (residues Tyr28–Asn34) in the structure of the mature enzyme (Figure 4, Table S3). During zymogen activation, proteolytic removal of the propeptide induces a global conformational change in which

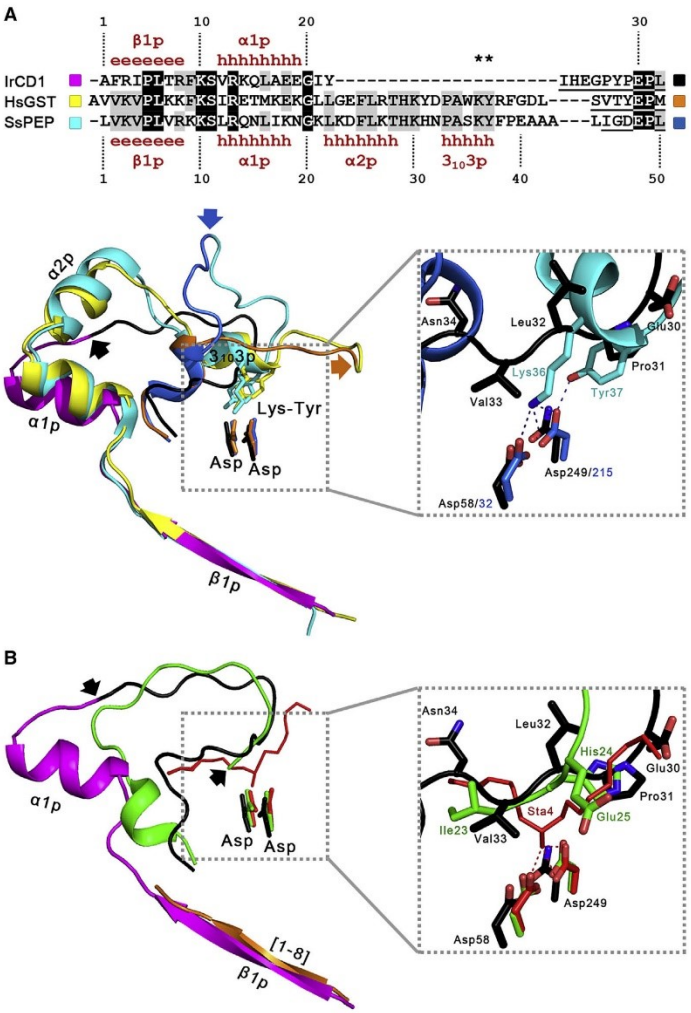


Figure 3. Structure of the IrCD1 Propeptide

(A) Comparison of the propeptides of IrCD1 and two archetypal pepsin-family aspartic peptidases, pig pepsin (SsPEP) and human gastricsin (HsGST). Structure-based sequence alignment shows the propeptides followed by the N termini of the mature enzymes (underlined). Identical and conserved residues are shaded black and gray, respectively. Residue numbering and secondary structure elements (in brown: e, β strand; h, helix) are shown for IrCD1 and pepsin; the anchor residues are marked with asterisks. Superposition of 3D structures is shown in cartoon representation. The propeptide and the following mature region are shown in magenta and black, respectively, for the IrCD1 zymogen; cyan and dark blue for pepsinogen; and yellow and orange for progastricsin (the color coding is indicated in the alignment). Arrows mark the positions of the cleavage site at the propeptide-enzyme core junction, which is processed to give rise to the mature enzymes. The secondary structure elements in the propeptides are labeled. The boxed area shows the interactions of the pepsinogen and progastricsin propeptides with two catalytic aspartate residues in the active site through the Lys-Tyr anchor residues. In the zoom view, the interacting residues (cyan and dark blue sticks) and hydrogen bonds (dashed lines) are shown for pepsinogen. Note that this interaction mechanism is absent in IrCD1 (black).

(B) Superposition of the propeptide of the IrCD1 zymogen, depicted as in (A), and the complex of the propeptide-derived inhibitor [1–8] (orange) with mature IrCD1 (green). The β1p segment of intact propeptide and its synthetic fragment [1–8] bind analogously to the exosite on the enzyme core. The arrows mark the cleavage site and the generated mature N terminus. The zoom view shows the mature N terminus of the [1–8] complex located in the vicinity of two catalytic aspartate residues in the active site. The substrate-analog inhibitor pepstatin (red backbone), which forms hydrogen bonds with the catalytic aspartates, is superimposed. Relevant residues are depicted as sticks. Uniprot and PDB accession numbers: IrCD1, A4GTA5; HsGST, P20142, 1HTR; and SsPEP, P00791, 3PSG.

the newly formed mature N terminus moves approximately 50 Å (Figure 2). It adopts a β strand secondary structure and binds in the position originally occupied by the N-terminal β strand of the propeptide (Figure 4). The N-terminal β strand segments in the zymogen and the mature enzyme share a high degree of structural similarity, which is also reflected in their sequence homology with conserved Pro-Leu residues as a dominant determinant (positions 5–6 and 31–32 in the propeptide and mature enzyme, respectively) (Figure 4D). Based on the interaction mode of these intramolecular ligands, we propose the term "exosite" to describe their binding site on the IrCD1 enzyme core located outside the active-site cleft. The exosite subsites are denoted $S1^E$ to $S8^E$ and the corresponding ligand residues $P1^E$ to $P8^E$; they are numbered according to the propeptide.

Furthermore, we performed sequence and structure alignments of the exosite ligands of IrCD1 with homologous pepsin-family peptidases (Figure 5). This analysis clearly showed that there is substantial similarity (1) among the propeptide N-terminal exosite ligands with a strong preference for Arg/Lys, Pro, Leu, and Arg/Lys in the $P3^E$, $P5^E$, $P6^E$, and $P8^E$ positions, respectively; and (2) among the mature N-terminal exosite ligands with a strong preference for Glu, Pro, Leu, and Asn in the $P4^E$, $P5^E$, $P6^E$, and $P8^E$ positions, respectively (Figure 5). Pro-Leu residues in $P5^E$–$P6^E$ are a common motif for all exosite ligands. These data indicate that the interaction of the exosite with its intramolecular ligands represents an evolutionarily conserved mechanism in pepsin-family peptidases.

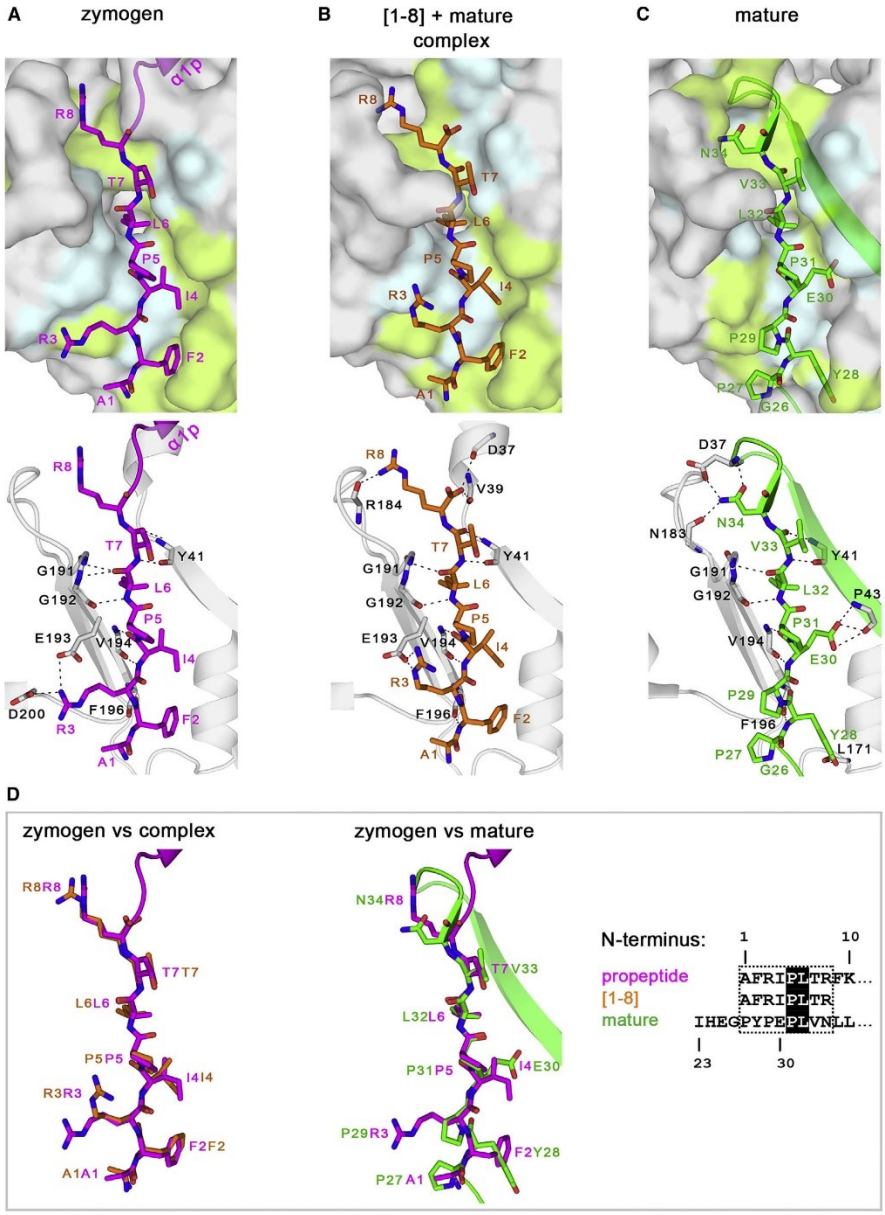


Figure 4. Binding of Natural Ligands to the Exosite on the IrCD1 Enzyme Core

(A–C) Interactions of three different ligands in the same exosite are compared: (A) the N-terminal region (residues 1–8) (magenta) of the intact propeptide in the structure of the IrCD1 zymogen, (B) the synthetic peptide derived from propeptide residues 1–8 (orange) in complex with mature IrCD1, and (C) the mature N terminus (residues 26–34) (green) in the structure of mature IrCD1. The ligands are in stick representation (heteroatoms are colored red and blue for oxygen and nitrogen, respectively) followed by a stretch in cartoon representation (colored as for ligands). In the upper panels, the enzyme core is shown as a gray

(legend continued on next page)

The active site of the IrCD1 zymogen is blocked by an Ω-shaped structure formed by the future mature N terminus (about 15 residues) and the upstream propeptide (particularly its α helix), which prevents substrate access to the active site. The residues Asp37, Asn34, Val33, and Leu32 interact with the substrate-binding subsites S3, S2, S1, and S3′, respectively (Figure 3B). The Ω-shaped structure is also stabilized by internal interactions, with Asn34 playing a central role. Importantly, there is no direct (or water-mediated) interaction of this structure with the catalytic Asp residues. This is a striking difference between the IrCD1 zymogen and zymogens of the classical gastric aspartic peptidases pepsin and gastricsin, in which the propeptide enters the catalytic site and interacts with both Asp residues through the conserved Lys-Tyr "anchor" residues (residues 36p–37p in pepsinogen propeptide numbering) (Figure 3A).

In summary, our structural analysis of the IrCD1 zymogen revealed an evolutionarily conserved exosite involved in the inactivation of pepsin-family zymogens by propeptides as tethered intramolecular ligands, as well as a non-conserved interaction with the active site.

The Exosite-Binding Segment of the Propeptide Inhibits the Mature Enzyme

Crystallographic analysis of the IrCD1 zymogen demonstrated that the propeptide and mature N-terminal region are responsible for inactivation of the active site. Next, we investigated whether these regions, in the form of synthetic peptides, can interact with and inhibit mature IrCD1. The peptide fragments were synthesized with capping groups (except for the native propeptide terminus) to mimic peptide linkages. The peptides were screened in a kinetic activity assay with mature IrCD1, and half-maximal inhibitory concentrations (IC_{50}) were determined. The mapping showed that the full-length propeptide (residues 1–22) exhibits strong inhibitory activity with an IC_{50} of ~13 nM, whereas the mature N-terminal region (residues 23–34) and a segment bridging both regions (peptides within residues 9–28) display no inhibitory activity (Figure 6A). The mode of inhibition of the synthetic propeptide determined by kinetic analysis was mixed-type, noncompetitive inhibition (Figure S4).

To identify the critical inhibitory region in the propeptide, we performed a deletion analysis using a series of truncated fragments. Gradual removal of 14 C-terminal residues from the propeptide did not substantially affect the inhibition (IC_{50} in the range of approximately 7–13 nM). Further trimming of one or two residues from the C-terminal side or at the N terminus, however, resulted in greatly increased IC_{50} values. This analysis demonstrated that the critical inhibitory region is located at the N terminus of the propeptide, spanning residues 1 to 8. The corresponding peptide [1–8] displayed an IC_{50} of ~11 nM in a kinetic analysis of IrCD1 inhibition (Figure 6A). Its binding to IrCD1 was also measured independently of enzyme activity using microscale thermophoresis, which yielded a K_d of ~3 nM, in good agreement with the IC_{50} value (Figure S5A).

The inhibitory potency of [1–8] was pH dependent (Figure S5B). We found that neutral pH (at and above pH 6.5) is required for effective formation of the enzyme-inhibitor complex. The preformed complex was stable at mildly acidic pH, leading to the strongest inhibition of IrCD1 at pH 5 (IC_{50} of ~11 nM); however, as the pH was further lowered, the inhibition effect gradually diminished, resulting in a three-order-of-magnitude higher IC_{50} value at pH 3.5. This pH-dependent behavior is reminiscent of the inactivity and activation of the zymogen at neutral and acidic pH values, respectively. Based on these observations, we employed a kinetic activity assay with a pH shift, including preincubation at pH 7 followed by fluorogenic substrate cleavage at pH 5 (Figure 6), to screen the propeptide-derived inhibitors against IrCD1.

Finally, we examined the selectivity of recognition of the N-terminal region of the propeptide by IrCD1 using propeptide sequences of related enzymes from the pepsin family (Figure 6B). We focused on propeptide fragments corresponding with residues 1–14 of the IrCD1 propeptide because there is significant sequence homology and structural similarity in this region among the pepsin-family propeptides (Figures 3A and 5). The synthetic peptides were evaluated for their inhibitory potency against mature IrCD1 in a kinetic activity assay; the obtained IC_{50} values spanned three orders of magnitude. The data indicated that the cross-inhibition reflects the degree of sequence homology and further highlighted the importance of exosite-binding residues for the interaction (Figure 6B).

Overall, mapping the IrCD1 propeptide identified the N-terminal segment that possesses an intrinsic inhibitory function and its minimal inhibitory motif. The corresponding synthetic peptide [1–8] is an efficient inhibitor of mature IrCD1 with a pH-dependent interaction. The inhibitory segment is recognized by IrCD1 with a substantial degree of selectivity over other pepsin-family propeptide sequences.

Structural Characterization of Allosteric Inhibition by a Propeptide-Derived Exosite Inhibitor

The synthetic peptide [1–8] derived from the N terminus of the IrCD1 propeptide is a nanomolar inhibitor of mature IrCD1. To analyze its inhibition mechanism, we determined the crystal structure of the complex of [1–8] with the mature enzyme at 1.81 Å resolution (Figure 2). This demonstrated that the inhibitor interacts with the propeptide-binding exosite on the enzyme core and that it has an analogous binding mode and β strand conformation as the parental N-terminal segment of the propeptide in the structure of the IrCD1 zymogen (Figure 4). The corresponding residues in both ligands occupy the same subsites of the exosite ($S1^E$ to $S8^E$), forming a similar network of hydrogen bonds and contacts (Table S3). A partial rearrangement of

surface with residues forming the interface hydrogen bonds in pale cyan and contact residues in lime green. In the lower panels, the residues of the enzyme core forming the interface hydrogen bonds (dashed lines) are shown as gray sticks. Ligand residues and the enzyme core residues are labeled in color-coded format. (D) Superposition of the ligands presented in (A–C): the IrCD1 zymogen versus mature IrCD1 (right), and the IrCD1 zymogen versus the complex of [1–8] with mature IrCD1 (left). Color coding of ligands is as above. The structure-based sequence alignment compares the N-terminal regions of the propeptide and the mature enzyme; the synthetic peptide [1–8] spans propeptide residues 1–8. The boxed area indicates the position of exosite ligands. Conserved residues are shaded black; residue numbering is according to the zymogen. Further details on ligand interactions are given in Table S3 and Figure S3.

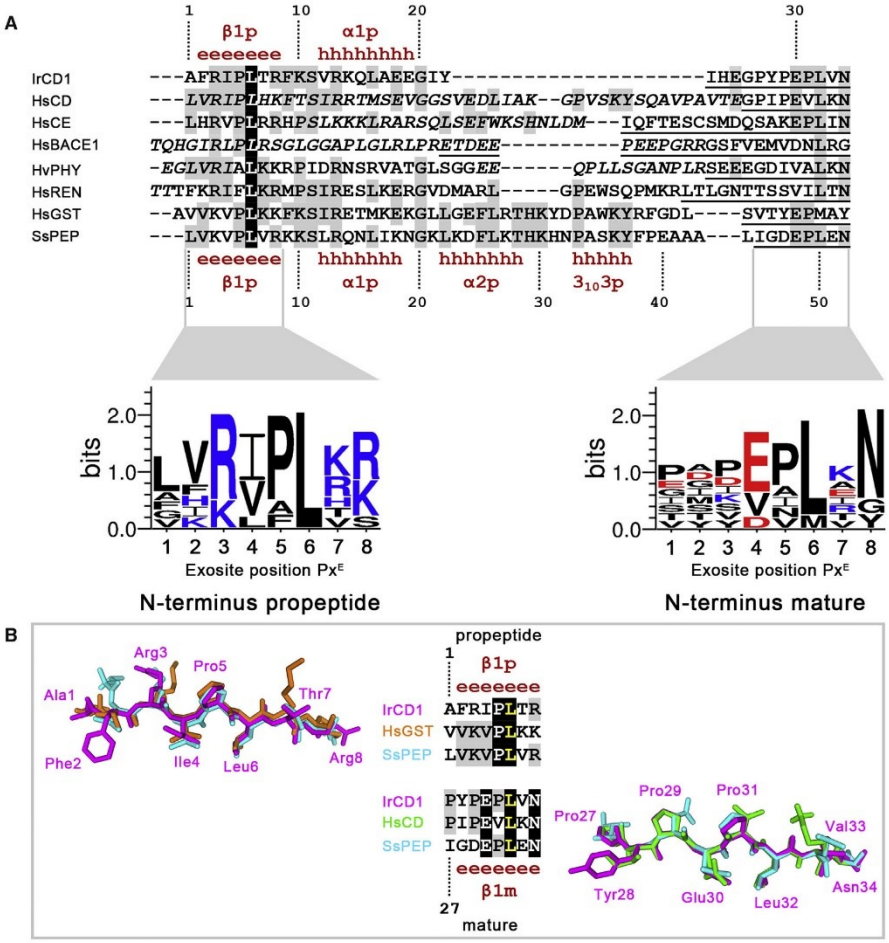


Figure 5. Homology of Intramolecular Exosite Ligands in Pepsin-Family Peptidases

(A) Structure-based sequence alignment of the propeptides and N termini of mature enzymes of pepsin-family aspartic peptidases including IrCD1, human cathepsin D (HsCD), human cathepsin E (HsCE), human β-secretase (HsBACE1), barley phytepsin (HvPhy), human renin (HsREN), human gastricsin (HsGST), and pig pepsin (SsPEP). Residues not resolved in the crystal structures of the zymogens are in italics; the mature N-terminal sequences are underlined. Identical and conserved residues are shaded black and gray, respectively. Residue numbering and secondary structure elements of the propeptides (in brown; e, β strand; h, helix) are shown for IrCD1 and pepsin. Sequence logo graphs shown below the alignment highlight residue conservation in two segments that interact with the same exosite on the enzyme core, as observed in structures of the zymogens (N-terminus propeptide) and mature enzymes (N-terminus mature). The $P1^E$ to $P8^E$ positions of residues of the exosite ligands are marked; note the highly conserved Leu6 position ($P6^E$). Logos were generated using WebLogo (Crooks et al., 2004); acidic and basic residues are colored red and blue, respectively.

(B) Structural homology of exosite ligands. Right: superposition of the N-terminal segment of the propeptide that binds the exosite in representative zymogens. Left: superposition of the mature N-terminal segment that binds the exosite in representative mature enzymes. Structures are shown in stick representation: IrCD1 (magenta), HsGST (orange), SsPEP (cyan), and HsCD (green). Middle: sequence alignment of superimposed structures; the invariant leucine residue (the $P6^E$ position) is yellow. The β strand conformation of both types of ligands is indicated. Residue numbering is according to the IrCD1 zymogen. Uniprot and PDB accession numbers: IrCD1, A4GTA5; HsCD, P07339, 1LYA; HsCE, P14091, 1TZD; HsBACE1, P56817, 3LPK; HvPhy, P42210, 1QDM; HsREN, P00797, 3VCM (the first three residues are not shown); HsGST, P20142, 1HTR; and SsPEP, A1YM35, 3PSG, 5PEP.

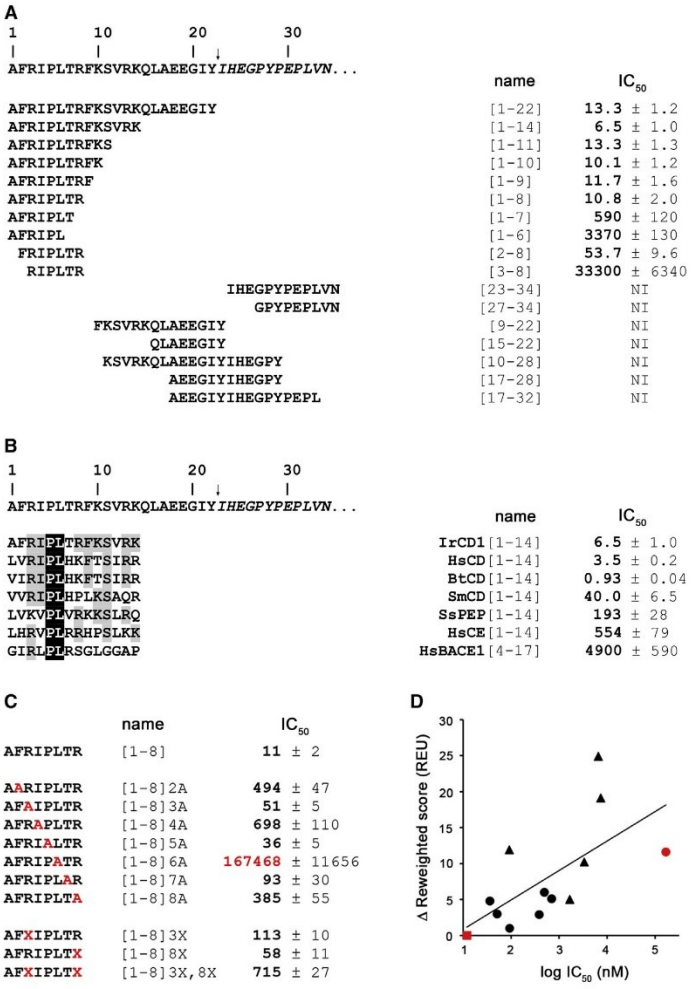


Figure 6. Inhibitory Activity of Synthetic Peptide Fragments Derived from Propeptides of IrCD1 and Its Homologs

(A) The amino acid sequence of the IrCD1 propeptide followed by the N-terminal sequence of mature IrCD1 (in italics) is shown in the upper line; the autoprocessing site is indicated with an arrow. Synthetic peptide fragments are coded according to their location in the sequence. The peptides were screened for their potential to inhibit the activity of mature IrCD1; half-maximal inhibitory concentrations (IC_{50}) were determined in a kinetic assay with an FRET-peptide substrate (the inhibitor complex was preformed at pH 7 and assayed at pH 5). Mean values ±SE are given for triplicate measurements; NI indicates no significant inhibition with 200 μM peptide.

(B) The synthetic fragment IrCD1 [1–14] derived from the IrCD1 propeptide (upper line) is compared with homologous fragments derived from the propeptides of representative pepsin-family proteases (HsCD, human cathepsin D; BtCD, bovine cathepsin D; SmCD, *Schistosoma mansoni* cathepsin D; SsPEP, pig pepsin; HsCE, human cathepsin E; HsBACE1, human β-secretase). Fragments are coded according to their location in the parental sequence. Residues identical to those of the IrCD1 fragment are shaded in gray; conserved residues are shaded in black. The peptides were screened for their potential to cross-inhibit the activity of mature IrCD1; IC_{50} values were determined as described above. Mean values ±SE are given for triplicate measurements.

(C) Inhibitory potency of synthetic derivatives of inhibitor [1–8], incorporating systematic substitutions with alanine and charge-changing substitutions of arginine with isosteric citrulline (denoted X). The substituted residues are shown in red. The peptides were screened for their ability to inhibit mature IrCD1, and IC_{50} values were determined (highlighted for critical substitution of Leu6). Mean values ±SE are given for triplicate measurements.

(D) Computational analysis of the binding energy of inhibitor [1–8] (red square) substituted with canonical amino acids highlights the dominant contribution of Leu6 and correlates well with inhibitory potency (R = 0.65, p value = 0.016). Systematic alanine scanning substitutions (circles, red for position 6) and five randomly selected hydrophobic substitutions at position 6 (triangles) are presented. Reweighted scores (in Rosetta energy units [REU]) were calculated using the Rosetta FlexPepDock protocol and compared with the wild-type inhibitor.

hydrogen bonds resulting from changes in the orientation of the Arg3 and Arg8 side chains occurred.

Comparing the structures of the unbound and inhibitor-bound mature enzymes, a global conformational change at the enzyme core is induced upon inhibitor [1–8] binding. The mature N terminus is displaced by the inhibitor from the exosite and moves by roughly 35 Å into the active site (Figure 2). The N-terminal stretch of 16 residues forms an Ω-shaped structure, which runs through the active-site cleft. Importantly, this feature clearly differs from the Ω-shaped structure that blocks the active site in the zymogen and is oriented in the reverse direction (Figure 3B). In the complex of [1–8], the displaced mature N terminus loses its β strand conformation and fills the substrate-binding subsites S1, S2′, and S3′ with the residues Ile23, His24, and Glu25, respectively. The N-terminal Ile23 forms indirect hydrogen bonding through a water molecule with the catalytic residues Asp58 and Asp249 (Figure S7A). The blocking of the active-site cleft is associated with local conformational changes in the tyrosine flap and polyproline loop moving in and out of the active site, respectively, compared with the unbound mature enzyme.

In conclusion, the structure of the IrCD1 complex with the propeptide-derived inhibitor [1–8] demonstrated that this inhibitor functions as an exosite ligand that allosterically regulates the active site. This represents a new mechanism for inhibition of

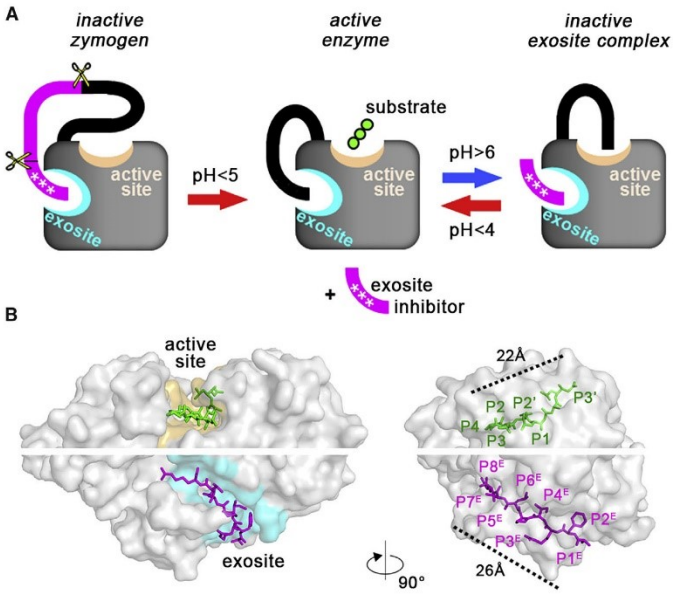


Figure 7. Central Role of the Exosite in Regulation of IrCD1

(A) A mechanistic model of pH-driven regulation of IrCD1 activity proposed in relation to IrCD1 function in the digestive vacuole of tick gut cells. The IrCD1 zymogen is inactivated through intramolecular autoinhibition by the propeptide (magenta), the N-terminal segment of which is bound to the exosite. The propeptide is autocatalytically removed upon primary acidification, generating the active mature enzyme together with a free fragment derived from the propeptide N terminus. The fragment acts as an allosteric exosite inhibitor when vacuolar pH increases due to extracellular cargo delivery. The fragment rebinds the exosite and induces conformational changes, leading to inhibition of the active site. The exosite inhibitor controls IrCD1 reversibly in a pH-dependent manner; the inhibition is reversed upon final acidification. This event activates IrCD1, which in turn initiates the proteolytic degradation pathway for ingested blood proteins that occurs optimally at acidic pH.

(B) Location of the exosite and the active site on the IrCD1 molecule shown in surface representation. The position of the exosite is indicated by the exosite inhibitor [1–8] (magenta sticks) with residue positions marked $P1^E$ to $P8^E$ (propeptide numbering). The position of the active site is indicated by the substrate-analog inhibitor pepstatin (green sticks); residue positions P3 to P3' (substrate numbering) are marked. IrCD1 residues interacting with the inhibitors are colored in the front view (left), and the size of the inhibitors is indicated in the side view (right; the enzyme surface is transparent). The model combines crystal structures of two inhibitor complexes.

mature aspartic peptidases by an intermolecular exosite-targeting inhibitor, which is clearly distinct from directly targeting the active site with substrate-mimicking inhibitors.

Structural Requirements for Binding the Exosite Inhibitor

Substituted derivatives of the octapeptide inhibitor [1–8] and chemical calculations of binding energies were employed to explore the structural requirements of the IrCD1 inhibitor-binding exosite. First, we performed systematic alanine scanning in inhibitor positions $P1^E$ to $P8^E$, in which each residue of the wild-type peptide was individually replaced with an Ala residue. The inhibitory potencies of the synthetic peptides were determined in a kinetic activity assay against mature IrCD1 (Figure 6C). All substitutions led to inhibitors with weaker potency. Most Ala substitutions resulted in IC_{50} values within one order of magnitude of the wild-type value, except for replacement of the Leu residue in position $P6^E$, which resulted in a dramatic, four-order-of-magnitude decrease in inhibition. We also substituted positively charged Arg residues in the inhibitor scaffold with isosteric citrulline containing an uncharged side chain. This had a moderate effect on the inhibition, pointing to the supportive role of side-chain hydrogen bonding and ionic interactions (formed by $P3^E$ and $P8^E$; Figure 4).

A computational alanine scan of the wild-type inhibitor [1–8] showed good correlation with the inhibition data, again highlighting Leu at the $P6^E$ position as the largest favorable contributor to ligand binding (Figure 6D). Furthermore, the residue specificity of the corresponding binding subsite, located in a deep hydrophobic pocket (Figure S7B), was computationally analyzed. Modeling of docking complexes with substitutions of the Leu residue revealed that all other proteinogenic amino acids are unfavorable in $P6^E$ (Table S4). Interestingly, methionine, which ranked second to leucine in this analysis, is the only rare mutation at the highly conserved Leu $P6^E$ position in natural intramolecular ligands of pepsin-family peptidases (Figure 5A). In summary, mapping of the structural requirements for binding the propeptide-derived inhibitor [1–8] to the IrCD1 exosite identified $P6^E$ as a "hotspot" position in the inhibitor scaffold.

DISCUSSION

In this work, we describe a dual regulatory function for the propeptide in pepsin-family aspartic peptidases and provide evidence for two different inhibition mechanisms, intramolecular and intermolecular, that employ an evolutionarily conserved propeptide-binding exosite. These findings result from an in-depth investigation of IrCD1, a parasite cathepsin D peptidase from the hard tick *I. ricinus*, at the molecular and structural levels.

Intramolecular autoinhibition controls the IrCD1 zymogen. The zymogen structure shows that the intact propeptide (and the adjacent mature region) fills the active site of IrCD1 to block substrate access. This interaction is guided by the N terminus of the propeptide located outside the active-site cleft, where it is tightly bound as an intrinsic tethered ligand to the exosite on the enzyme core. Sequence and structural comparison revealed that the 22-residue propeptide of the IrCD1 zymogen represents a minimalist version of the propeptides of other pepsin-family

peptidases, which can be up to twice as long. Importantly, we demonstrate that the exosite-binding segment at the N terminus is evolutionarily conserved in pepsin-family propeptides, as is its mode of interaction with the exosite. In contrast, there is high variability in the C-terminal part of the propeptide directed into the active-site cleft. In mammalian gastric aspartic peptidases and cathepsin D, this part is extended, and two anchor residues directly interact with the catalytic aspartates (Figure 3B) (James and Sielecki, 1986; Masa et al., 2006; Moore et al., 1995). In the shorter propeptides of mammalian renin and plant phytepsin, the C-terminal anchor segment is missing, but its interaction with the catalytic aspartates is mimicked by residues from the adjacent mature region (Figure S6) (Kervinen et al., 1999; Morales et al., 2012). Further reduction of the propeptide length and structural simplification, as seen in the IrCD1 zymogen, results in loss of direct interaction with the catalytic residues. This comparison underscores the critical role played by the N-terminal part of the propeptide pinned to the conserved exosite, while steric blockage of the active site by the downstream part of the propeptide can be achieved in multiple ways.

We discovered a novel regulatory mechanism mediated by the exosite leading to intermolecular inhibition of mature IrCD1. During autoactivation of the IrCD1 zymogen at acidic pH, the propeptide is cleaved off and fragmented, generating a short oligopeptide fragment from the N terminus of the propeptide. We found that this fragment acts as a potent inhibitor of mature IrCD1. The crystal structure of the inhibitory fragment in complex with mature IrCD1 revealed that it functions as an exosite inhibitor through a unique allosteric mechanism. The binding of the inhibitor to the exosite mimics that identified for the intact propeptide of the IrCD1 zymogen; however, the blockage of the enzyme active site is strikingly different. Exosite occupation by the inhibitor induces a large conformational change in the mature N terminus, which is displaced (by approximately 35 Å) into the active site, forming a docking interaction with the catalytic aspartates via water-mediated hydrogen bonding (Figure S7A). This represents a novel mode of structural occlusion of the active site, differing from that observed in the zymogens of aspartic peptidases or activation intermediates containing the cleaved, but not yet liberated, propeptide (Khan et al., 1997; Ostermann et al., 2004).

The potency of the exosite inhibitor of IrCD1 was pH dependent. The inhibitor binds to IrCD1 at pH values above 6, forming a stable complex. The IrCD1 activity of the complex is effectively inhibited above pH ~4.0; i.e., above the enzyme's pH optimum (IrCD1 is active at acidic pH up to pH 5.5, with a pH optimum at ~4.0; Sojka et al., 2012). This behavior can be explained by the accessibility of the exosite at different pH values (Figure S5B). The exosite preferentially binds the mature N terminus at acidic pH (in mature IrCD1 and other pepsin-family peptidases) and the propeptide N terminus at neutral pH (in the zymogens). Thus, for the mature enzyme, an increase in pH toward neutral will destabilize the mature N terminus (as observed in human cathepsin D; Lee et al., 1998) and open the exosite to the inhibitory ligand.

We suggest that this pH-modulated exosite inhibition of IrCD1 is biologically relevant. IrCD1 is localized in the digestive vacuoles of tick gut cells (Sojka et al., 2012), which are functional analogs of mammalian endolysosomes; their biogenesis is associated with continual pH change spanning more than 2.5 pH units (Lara et al., 2005; Luzio et al., 2014; Sojka et al., 2013). We propose the following scenario for IrCD1 regulation (Figure 7). First, in primary acidified lysosomes, the IrCD1 zymogen is autoprocessed to mature IrCD1, generating the free exosite inhibitor. Fusion of these small vesicles with the large endocytic vesicles (delivering blood protein cargo at around neutral pH) increases pH, enabling the exosite inhibitor to bind and inactivate IrCD1.

The resulting digestive vacuoles are then gradually acidified by vacuolar-type ATPases (Mindell, 2012) to become competent for acidic proteolysis; however, IrCD1 remains inactive until it reaches a deep acidic status when the exosite inhibitor unbinds. Our previous work showed that IrCD1 is the critical enzyme responsible for initiation of the degradation of host hemoglobin at low acidic pH, which is then followed by the action of other downstream peptidases in the hemoglobinolytic pathway (Horn et al., 2009). Hence, the inhibition mechanism targeting IrCD1 and driven by endolysosomal pH helps precisely control triggering of blood protein proteolysis in the tick digestive organelle.

We show that the structure of the exosite and its recognition motif in the propeptide are highly conserved in pepsin-family peptidases (Figure 5). This is further supported by functional data on cross-recognition of the exosite-binding motif derived from various family members (Figure 6B). These findings, together with previous reports on the inhibition properties of processed propeptides (for references, see Masa et al., 2006), suggest that the exosite allosteric inhibition mechanism might be involved in the regulation of not only IrCD1 but also other members of the pepsin family. Further research will be needed to evaluate the role of allosteric inhibition in physiologically relevant proteolytic processes governed by these enzymes. In complex systems, a secondary regulatory level based on reversible inhibition of the mature enzyme may provide an advantage over simple regulation by zymogen activation.

Our work provides a basis for the rational design of a novel class of allosteric inhibitors of pepsin-family peptidases based on the exosite ligand scaffold. Analysis of the structural requirements of the IrCD1 exosite inhibitor identified the minimal binding motif and the hotspot position critical for the inhibitory interaction. These structure-inhibition relationships and validated computational prediction method allow for the engineering of unnatural derivatives of the scaffold with tailored properties, including inhibitor potency and selectivity. Such regulators that could sense pH changes under specific pathological conditions would be a novel tool to target medicinally relevant enzymes from the pepsin family.

SIGNIFICANCE

Aspartic peptidases (proteases) of the pepsin family are broadly distributed and critically involved in numerous proteolytic processes and pathologies, including Alzheimer disease, cancer, and malaria. Due to their importance, aspartic peptidases are tightly controlled in biological systems. However, little is known about the molecules that regulate their physiological functions, and, surprisingly, no selective endogenous inhibitors have been identified to date for family members in higher organisms. Here, we investigated at the

molecular and structural levels the aspartic peptidase IrCD1 from the tick *Ixodes ricinus*, a blood-feeding parasite that transmits Lyme disease and tick-borne encephalitis. Our findings revealed a fundamentally new regulatory mechanism in aspartic peptidases involving an oligopeptide inhibitor released from the peptidase zymogen during autocatalytic processing. This inhibitor acts allosterically by binding to an exosite on the parental peptidase. The exosite is evolutionarily conserved, suggesting that the molecular pathway that generates the selective endogenous inhibitor is general for pepsin-family aspartic peptidases. Furthermore, the study provides a basis for rational design of a novel class of allosteric inhibitors targeting the exosites of various medically important aspartic peptidases.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Mutagenesis and Production of IrCD1 Zymogen
  - Preparation of Mature Enzyme and Complexes
  - Protein Crystallization and Data Collection
  - Structure Determination, Refinement, and Analysis
  - IrCD1 Activity and Inhibition Assay
  - Analysis of IrCD1 Zymogen Processing
  - Synthesis and Processing of Propeptide Fragments
  - Microscale Thermophoresis (MST)
  - *In Silico* Modeling of Peptide-IrCD1 Complexes
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and four tables and can be found with this article online at https://doi.org/10.1016/j.chembiol.2018.01.001.

ACKNOWLEDGMENTS

This work was supported by grants 15-18929S and 13-11043S from the Czech Science Foundation, project InterBioMed LO1302 from the Ministry of Education of the Czech Republic, Gilead Sciences and IOCB Research Center, and institutional project RVO 61388963. D.S. and P.K. were supported by grant 14-33693S from the CSF and institutional project RVO 60077344. N.A. and O.S.-F. were supported by grant 2015207 from the USA-Israel Binational Science Foundation. We thank Martin Hubálek and Jana Březinová for mass spectrometry, Mirka Blechová and Martin Hradilek for peptide synthesis, Pavlína Řezáčová for crystallography expertise, Jana Pytelková for biochemical experiments, Zdeněk Voburka for Edman sequencing (Institute of Organic Chemistry and Biochemistry, C.A.S., Prague, Czech Republic), and Hillary Hoffman for critical proofreading of the manuscript.

AUTHOR CONTRIBUTIONS

I.H. and R.H. performed biochemical and enzymological experiments. I.H. prepared and crystallized proteins, and J.B. solved crystal structures. L.M. prepared and analyzed selected peptides. D.S. and P.K. provided expression constructs and contributed to biological interpretation. N.A., J.V., and O.S.-F. conceived, performed, and interpreted computational experiments. M.H. contributed to manuscript writing and data interpretation. M.M. conceived and supervised the study and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 5, 2017
Revised: December 4, 2017
Accepted: December 28, 2017
Published: January 25, 2018

SUPPORTING CITATIONS

The following reference appears in the Supplemental Information: Conner (1989).

REFERENCES

Alam, N., and Schueler-Furman, O. (2017). Modeling peptide-protein structure and binding using Monte Carlo sampling approaches: Rosetta FlexPepDock and FlexPepBind. Methods Mol. Biol. *1561*, 139–169.

Alam, N., Zimmerman, L., Wolfson, N.A., Joseph, C.G., Fierke, C.A., and Schueler-Furman, O. (2016). Structure-based identification of HDAC8 non-histone substrates. Structure *24*, 458–468.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Collaborative Computational Project, Number 4 (1994). The CCP4 suite - programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Conner, G.E. (1989). Isolation of procathepsin D from mature cathepsin D by pepstatin affinity chromatography. Autocatalytic proteolysis of the zymogen form of the enzyme. Biochem. J. *263*, 601–604.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res. *14*, 1188–1190.

Doleckova-Maresova, L., Pavlik, M., Horn, M., and Mares, M. (2005). De novo design of alpha-amylase inhibitor: a small linear mimetic of macromolecular proteinaceous ligands. Chem. Biol. *12*, 1349–1357.

Dunn, B.M. (2002). Structure and mechanism of the pepsin-like family of aspartic peptidases. Chem. Rev. *102*, 4431–4458.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Franta, Z., Frantova, H., Konvickova, J., Horn, M., Sojka, D., Mares, M., and Kopacek, P. (2010). Dynamics of digestive proteolytic system during blood feeding of the hard tick *Ixodes ricinus*. Parasit. Vectors *3*, 119.

Horn, M., Jilkova, A., Vondrasek, J., Maresova, L., Caffrey, C.R., and Mares, M. (2011). Mapping the pro-peptide of the *Schistosoma mansoni* cathepsin B1 drug target: modulation of inhibition by heparin and design of mimetic inhibitors. ACS Chem. Biol. *6*, 609–617.

Horn, M., Nussbaumerova, M., Sanda, M., Kovarova, Z., Srba, J., Franta, Z., Sojka, D., Bogyo, M., Caffrey, C.R., Kopacek, P., et al. (2009). Hemoglobin digestion in blood-feeding ticks: mapping a multipeptidase pathway by functional proteomics. Chem. Biol. *16*, 1053–1063.

James, M.N.G., and Sielecki, A.R. (1986). Molecular-structure of an aspartic proteinase zymogen, porcine pepsinogen, at 1.8 A resolution. Nature *319*, 33–38.

Jilkova, A., Horn, M., Rezacova, P., Maresova, L., Fajtova, P., Brynda, J., Vondrasek, J., McKerrow, J.H., Caffrey, C.R., and Mares, M. (2014). Activation route of the *Schistosoma mansoni* cathepsin B1 drug target: structural map with a glycosaminoglycan switch. Structure *22*, 1786–1798.

Kervinen, J., Tobin, G.J., Costa, J., Waugh, D.S., Wlodawer, A., and Zdanov, A. (1999). Crystal structure of plant aspartic proteinase prophytepsin: inactivation and vacuolar targeting. EMBO J. *18*, 3947–3955.

Khan, A.R., Cherney, M.M., Tarasova, N.I., and James, M.N. (1997). Structural characterization of activation 'intermediate 2' on the pathway to human gastricsin. Nat. Struct. Biol. *4*, 1010–1015.

Khan, A.R., Khazanovich-Bernstein, N., Bergmann, E.M., and James, M.N.G. (1999). Structural aspects of activation pathways of aspartic protease zymogens and viral 3C protease precursors. Proc. Natl. Acad. Sci. USA *96*, 10968–10975.

Kuhlman, B., and Baker, D. (2000). Native protein sequences are close to optimal for their structures. Proc. Natl. Acad. Sci. USA *97*, 10383–10388.

Lara, F.A., Lins, U., Bechara, G.H., and Oliveira, P.L. (2005). Tracing heme in a living cell: hemoglobin degradation and heme traffic in digest cells of the cattle tick *Boophilus microplus*. J. Exp. Biol. *208*, 3093–3101.

Leaver-Fay, A., O'Meara, M.J., Tyka, M., Jacak, R., Song, Y., Kellogg, E.H., Thompson, J., Davis, I.W., Pache, R.A., Lyskov, S., et al. (2013). Scientific benchmarks for guiding macromolecular energy function improvement. Methods Enzymol. *523*, 109–143.

Lee, A.Y., Gulnik, S.V., and Erickson, J.W. (1998). Conformational switching in an aspartic proteinase. Nat. Struct. Biol. *5*, 866–871.

Lee, J., Shin, M.K., Ryu, D.K., Kim, S., and Ryu, W.S. (2010). Insertion and deletion mutagenesis by overlap extension PCR. Methods Mol. Biol. *634*, 137–146.

Lenarcic, B., and Turk, V. (1999). Thyroglobulin type-1 domains in equistatin inhibit both papain-like cysteine proteinases and cathepsin D. J. Biol. Chem. *274*, 563–566.

Li, M., Phylip, L.H., Lees, W.E., Winther, J.R., Dunn, B.M., Wlodawer, A., Kay, J., and Gustchina, A. (2000). The aspartic proteinase from *Saccharomyces cerevisiae* folds its own inhibitor into a helix. Nat. Struct. Biol. *7*, 113–117.

London, N., Lamphear, C.L., Hougland, J.L., Fierke, C.A., and Schueler-Furman, O. (2011). Identification of a novel class of farnesylation targets by structure-based modeling of binding specificity. PLoS Comput. Biol. *7*, e1002170.

Luzio, J.P., Hackmann, Y., Dieckmann, N.M., and Griffiths, G.M. (2014). The biogenesis of lysosomes and lysosome-related organelles. Cold Spring Harb. Perspect. Biol. *6*, a016840.

Mares, M., Meloun, B., Pavlik, M., Kostka, V., and Baudys, M. (1989). Primary structure of cathepsin D inhibitor from potatoes and its structure relationship to soybean trypsin inhibitor family. FEBS Lett. *251*, 94–98.

Masa, M., Maresova, L., Vondrasek, J., Horn, M., Jezek, J., and Mares, M. (2006). Cathepsin D propeptide: mechanism and regulation of its interaction with the catalytic core. Biochemistry *45*, 15474–15482.

Mindell, J.A. (2012). Lysosomal acidification mechanisms. Annu. Rev. Physiol. *74*, 69–86.

Moore, S.A., Sielecki, A.R., Chernaia, M.M., Tarasova, N.I., and James, M.N.G. (1995). Crystal and molecular-structures of human progastricsin at 1.62 angstrom resolution. J. Mol. Biol. *247*, 466–485.

Morales, R., Watier, Y., and Bocskei, Z. (2012). Human prorenin structure sheds light on a novel mechanism of its autoinhibition and on its non-proteolytic activation by the (pro)renin receptor. J. Mol. Biol. *421*, 100–111.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Ostermann, N., Gerhartz, B., Worpenberg, S., Trappe, J., and Eder, J. (2004). Crystal structure of an activation intermediate of cathepsin E. J. Mol. Biol. *342*, 889–899.

Raveh, B., London, N., and Schueler-Furman, O. (2010). Sub-angstrom modeling of complexes between flexible peptides and globular proteins. Proteins *78*, 2029–2040.

Raveh, B., London, N., Zimmerman, L., and Schueler-Furman, O. (2011). Rosetta FlexPepDock ab-initio: simultaneous folding, docking and refinement of peptides onto their receptors. PLoS One *6*, e18934.

Rawlings, N.D., and Barrett, A.J. (2013). Introduction: aspartic and glutamic peptidases and their clans. In Handbook of Proteolytic Enzymes, Third Edition, *Vols. 1 and 2*, N.D. Rawlings and G. Salvesen, eds. (Elsevier), pp. 3–19.

Richter, C., Tanaka, T., and Yada, R.Y. (1998). Mechanism of activation of the gastric aspartic proteinases: pepsinogen, progastricsin and prochymosin. Biochem. J. *335*, 481–490.

Sojka, D., Franta, Z., Frantova, H., Bartosova, P., Horn, M., Vachova, J., O'Donoghue, A.J., Eroy-Reveles, A.A., Craik, C.S., Knudsen, G.M., et al. (2012). Characterization of gut-associated cathepsin D hemoglobinase from tick *Ixodes ricinus* (IrCD1). J. Biol. Chem. *287*, 21152–21163.

Sojka, D., Franta, Z., Horn, M., Caffrey, C.R., Mareš, M., and Kopáček, P. (2013). New insights into the machinery of blood digestion by ticks. Trends Parasitol. *29*, 276–285.

Vagin, A., and Teplyakov, A. (2000). An approach to multi-copy search in molecular replacement. Acta Crystallogr. D Biol. Crystallogr. *56*, 1622–1624.

Zebrakovska, I., Masa, M., Srp, J., Horn, M., Vavrova, K., and Mares, M. (2011). Complex modulation of peptidolytic activity of cathepsin D by sphingolipids. Biochim. Biophys. Acta *1811*, 1097–1104.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
E. coli TOP10	Invitrogen	Cat# C404010
E. coli BL21(DE3)	Invitrogen	Cat# C600003
Biological Samples		
Plasmid pET101/D-TOPO	Invitrogen	Cat# K10101
Chemicals, Peptides, and Recombinant Proteins		
Recombinant IrCD1	This paper	N/A
Recombinant IrCD1(D249N)	This paper	N/A
Q Sepharose HP HiLoad 16/10	GE Healthcare	Cat# 17-1064-01
Mono Q 5/50 GL	GE Healthcare	Cat# 17-5166-01
Mono S 5/50 GL	GE Healthcare	Cat# 17-5168-01
Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu	Laboratory of M. Mareš	Masa et al., 2006
Pepstatin A	Applichem	Cat# A2205
Dextran sulfate	Sigma-Aldrich	Cat# D6924
Heparin	Sigma-Aldrich	Cat# H3393
Bovine hemoglobin	Sigma-Aldrich	Cat# H2500
Potato cathepsin D inhibitor	Laboratory of M. Mareš	N/A
Synthetic propeptide-derived fragments	This paper	N/A
Synthetic derivatives of inhibitor [1-8]	This paper	N/A
Monolith NT-647-NHS	NanoTemper Technologies	Cat# L001
Deposited Data		
Crystal structure of mature IrCD1	This paper	PDB: 5N71
Crystal structure of IrCD1 zymogen	This paper	PDB: 5N7N
Crystal structure of IrCD1-[1-8] complex	This paper	PDB: 5N70
Crystal structure of IrCD1-pepstatin complex	This paper	PDB: 5N7Q
Oligonucleotides		
IrCD1-F: CACCATGGCTTTCAGGATCCCGCTCACG	This paper	N/A
IrCD1-R: CTAATGATGATGATGATGAT GGGCAACTTCCGCGAAGCC	This paper	N/A
IrCD1-M1: TGCGAAGCCATTGCGAACACGG	This paper	N/A
IrCD1-M2: CCGTGTTCGCAATGGCTTCGCA	This paper	N/A
Recombinant DNA		
Plasmid pET101-IrCD1	This paper	N/A
Plasmid pET101-IrCD1(D249N)	This paper	N/A
Software and Algorithms		
MOLREP	Vagin and Teplyakov, 2000	http://www.ccp4.ac.uk/html/molrep.html
REFMAC 5.2	Murshudov et al., 1997	http://www.ccp4.ac.uk/html/refmac5.html
Coot	Emsley and Cowtan, 2004	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
MolProbity	Chen et al., 2010	http://molprobity.biochem.duke.edu
PyMOL	Schrödinger Inc. NY	http://www.pymol.org
Rosetta Software Suite	RosettaCommons	https://www.rosettacommons.org
Grafit Data Analysis Software	Erithacus Software	http://www.erithacus.com
NT Analysis software 1.5.41	NanoTemper Technologies	N/A
Other		
Monolith NT.115 instrument	NanoTemper Technologies	N/A

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Michael Mareš (mares@uochb.cas.cz).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

E. coli TOP10 was used as a host for general cloning. *E. coli* BL21(DE3) was used for heterologous protein expression; cultures were grown in LB medium containing 50 μg ml^{-1} carbenicillin and after induction with 0.5 mM IPTG cultivated for 6 h at 37°C (Sojka et al., 2012).

METHOD DETAILS

Mutagenesis and Production of IrCD1 Zymogen

IrCD1 zymogen with a C-terminal His_6-tag was prepared in an *Escherichia coli* expression system using the pET101/D-TOPO expression vector (Invitrogen) and the IrCD1-F/IrCD1-R cloning primer pair. The active site mutant zymogen ($IrCD1^{as}$ zymogen) with a mutated catalytic residue (Asp249Asn) was constructed by site-directed mutagenesis. Mutagenesis was performed using the overlap extension PCR method (Lee et al., 2010) with the cloning primers and a pair of mutagenic primers IrCD1-M1/IrCD1-M2. The sequence-verified constructs were transformed into BL21(DE3) *E. coli* (Invitrogen), and the expressed proteins were isolated from inclusion bodies by chelating chromatography as previously described (Sojka et al., 2012). The denatured proteins were refolded using the following protocol: After reduction for 1 h at 26 °C in 50 mM Tris-HCl pH 8.0, 50 mM reduced glutathione, 8 M urea, the solution was diluted 10-fold to a concentration of 20 μg ml^{-1} into 50 mM Tris-HCl pH 8.0, 10% glycerol, 0.05% PEG 1500. The mixture was stirred for 0.5 h at 26 °C and incubated for 72 h at 4 °C. The refolded zymogens were purified by FPLC on a Q Sepharose HiLoad 16/10 column (GE Healthcare) as previously described (Sojka et al., 2012) followed by a chromatography on a Mono Q 5/50 GL column (GE Healthcare) equilibrated in 20 mM Bis-Tris pH 6.5 and eluted using a linear gradient of 0-1 M NaCl. IrCD1 zymogen with N-terminal His_6-tag was prepared using the *E. coli* expression system as previously described (Sojka et al., 2012). Purified IrCD1 zymogens were buffer exchanged into 10 mM Bis-Tris pH 6.5 and concentrated to 10 mg ml^{-1} using Amicon Ultracel-10K filters (Millipore).

Preparation of Mature Enzyme and Complexes

To obtain mature IrCD1, the N-terminally-tagged zymogen was autoactivated at pH 4.0 and purified on a Mono S 5/50 GL column (GE Healthcare) (Sojka et al., 2012). To prepare the active site-mutated mature $IrCD1^{as}$, the $IrCD^{as}$ zymogen was trans-activated by incubation with active mature IrCD1 (200:1 molar ratio) in 50 mM sodium acetate pH 4.0 at 26 °C and purified using a Mono S HR 5/5 column as above. Both types of mature enzymes were buffer exchanged into 10 mM Bis-Tris pH 6.5 and concentrated to 6 mg ml^{-1}. Inhibitor complexes of mature IrCD1 (5 mg ml^{-1}) were prepared by incubation with a 2-fold molar excess of pepstatin in 10 mM sodium formate pH 4.0 or a 3-fold molar excess of peptide [1-8] in 10 mM Tris-HCl pH 7.0.

Protein Crystallization and Data Collection

Crystals were grown by vapor-diffusion at 18°C in hanging drops containing a 1:1 ratio of protein to reservoir solution. Drops were equilibrated over 1 ml reservoir solution: 0.2 M ammonium sulfate, 0.1 M sodium acetate pH 4.6, 30% PEG MME 2000 for mature $IrCD1^{as}$; 140 mM sodium chloride, 70 mM Bis-Tris pH 5.5, 17.5% PEG 3350, 1 mM dithiothreitol for the IrCD1-pepstatin complex; 0.2 M sodium acetate, 0.1 M sodium cacodylate pH 6.5, 30% PEG 8000 for the IrCD1-[1-8] complex; and 85 mM Tris-HCl pH 8.0, 1.7 M ammonium sulfate, 1 mM dithiothreitol for $IrCD1^{as}$ zymogen. The protein concentrations of the stock solutions were 6, 5, and 10 mg ml^{-1} for mature $IrCD1^{as}$, IrCD1 complexes, and $IrCD1^{as}$ zymogen, respectively. The crystals of zymogen and mature $IrCD^{as}$ were flash cooled by plunging into liquid nitrogen without cryoprotectant. The crystals of IrCD1 in complex with [1-8] or pepstatin were cryoprotected in mother liquor supplemented with 25% glycerol or 30% PEG 3350, respectively, before flash cooling. Diffraction data from crystals of mature $IrCD1^{as}$ and IrCD1 complexes were collected at 100 K at beamlines MX14.1 and MX14.2, operated by the Helmholtz-Zentrum Berlin at the Bessy II electron storage ring, Berlin, Germany; data from crystals of $IrCD1^{as}$ zymogen were collected at ESRF beamline 14.4 in Grenoble, France. The crystal parameters and data collection statistics are given in Table S1.

Structure Determination, Refinement, and Analysis

The structure of mature IrCD1 in complex with pepstatin was determined by molecular replacement using human cathepsin D (PDB code 1LYB), which shares 52% sequence identity with IrCD1, as a model. The complex crystallized in the monoclinic space group $P2_1$ with two molecules in the asymmetric unit and a solvent content of 41%. The structure was refined using data to a resolution of 1.46 Å. Both molecules in the asymmetric unit were very similar, with a root mean square deviation (r.m.s.d.) for superposition of the two protein backbones of 0.63 Å. Minor structural changes were localized in the surface-exposed loops, namely in residues 231-235, 273-278, and 311-320. One and two N-terminal residues could not be located in the electron density map of molecules A and B, respectively.

The structure of uncomplexed mature IrCD^{as} was determined by molecular replacement using a model derived from the mature IrCD1-pepstatin complex. The enzyme crystallized in the tetragonal space group $P4_12_12$ with one single chain molecule in the asymmetric unit and a solvent content of 59%. The structure was refined using data to a resolution of 1.88 Å. Two N-terminal residues and the C-terminal His-tag could not be located in the electron density map and were thus not included in the final model.

The structure of the IrCD1^{as} zymogen was determined by molecular replacement using a model derived from the mature IrCD1-pepstatin complex. The zymogen crystallized in the monoclinic space group $P2_1$ with four molecules in the asymmetric unit (A to D), and the solvent content was 59%. The structure was refined to a resolution of 2.30 Å. The four molecules present in the asymmetric unit were very similar, with r.m.s.d. for the superposition of the protein backbones ranging from 0.27 to 0.36 Å. The N-terminal residue of molecule D and the C-terminal His tag of molecules A-D could not be located in the electron density maps and were not included in the final model.

The structure of the complex of mature IrCD1 with the octapeptide [1-8] was determined by molecular replacement using a model derived from the mature IrCD1-pepstatin complex. The complex crystallized in the tetragonal space group $P4_1$ with one molecule in the asymmetric unit and a solvent content of 44%. The structure was refined using data to a resolution of 1.81 Å. The electron density used for modeling of the ligand [1-8] was of excellent quality (Figure S3).

Molecular replacement was performed with the program MOLREP (Vagin and Teplyakov, 2000). Model refinement was carried out using the program REFMAC 5.2 (Murshudov et al., 1997) from the CCP4 package (Collaborative Computational Project, Number 4, 1994), interspersed with manual adjustments using Coot (Emsley and Cowtan, 2004). The quality of the final model was validated with MolProbity (Chen et al., 2010). The final refinement statistics are given in Table S1. All figures showing structural representations were prepared with the program PyMOL (Schrodinger LLC).

IrCD1 Activity and Inhibition Assay

The enzymatic activity of IrCD1 was measured in a continuous kinetic assay using the FRET substrate Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu (Masa et al., 2006). The assay was performed in 96-well microplates with a total reaction volume of 100 μl at 37 °C. The assay mixture contained 1.5 nM mature IrCD1 and 25 μM substrate in 0.1 M sodium acetate pH 4.0, 0.1% PEG 1500. The kinetics of product release were continuously monitored in an Infinite M1000 microplate reader (Tecan) at excitation and emission wavelengths of 330 and 420 nm, respectively. For inhibition measurements, 1.5 nM mature IrCD1 was preincubated with 0-200 μM synthetic peptide in 30 mM Tris-HCl pH 7.0, 0.1% PEG 1500 for 10 min at 37 °C, followed by addition of 25 μM substrate and 0.2 M sodium acetate pH 5.0, 0.1% PEG 1500. IC_{50} values were determined from residual velocities using dose-response plots (v_i/v_o vs. [I]); nonlinear regressions were fitted using Grafit software (Erithacus). Where indicated, the inhibition assay was modified as follows: (i) preincubation was performed at pH values ranging from 5.0 to 6.5 or (ii) the measurement was performed at pH values ranging from 3.0 to 4.5 (sodium acetate pH 3.0-5.5 and MES pH 6.0-6.5 were used as buffers). The inhibition mode was determined using an analogous activity assay with substrate concentration 7.5-30 μM; the initial velocities of product release were interpreted by Lineweaver-Burk plot. All experiments were performed in triplicate. The concentration of mature IrCD1 was determined by active site titration with pepstatin (Franta et al., 2010); the peptide solutions were quantified by amino acid analysis.

Analysis of IrCD1 Zymogen Processing

The C-terminally-tagged IrCD1 zymogen (15 nM) was autoactivated by incubation at 26 °C in 50 mM sodium acetate pH 4.0, 0.1% PEG 1500. Aliquots (10 μl) were taken from the reaction mixture at different time intervals, and the activity of mature IrCD1 was measured in a kinetic assay at pH 4.0. Where indicated, incubation was performed in the presence of 0.5 M NaCl, and 1 μg ml^{-1} dextran sulfate (D6924, Sigma) or heparin (H3393, Sigma). For pH profiling, the autoactivation was performed in 50 mM citrate-phosphate buffer pH 3.0-7.0, 0.1% PEG 1500 and assayed at pH 4.0. For imaging of zymogen conversion, the IrCD1 zymogen (0.5 μM) was autoactivated at 26 °C in 50 mM sodium acetate pH 4.0, 0.1% PEG 1500. Aliquots (20 μl) were taken from the reaction mixture at different time intervals, and proteins were acetone precipitated, resolved by Laemmli SDS-PAGE, and stained with silver nitrate. Where indicated, the incubation was performed in the presence of bovine hemoglobin (1.5 μM), pepstatin (2.5 μM), or PDI, a potato cathepsin D inhibitor (5 μM)(Mares et al., 1989). In the concentration dependence experiment, 0.05 μM and 5 μM IrCD1 zymogen were autoactivated, and 200 μl and 2 μl aliquots, respectively, were analyzed by Laemmli SDS-PAGE. The IrCD1^{as} zymogen (30 μM) was processed *in trans* by mature IrCD1 (150 nM) at 26 °C for 48 and 96 h in 50 mM sodium acetate pH 4.0, 0.1% PEG 1500. The reaction mixture was analyzed by Laemmli SDS-PAGE.

Processing products generated during autoactivation of the IrCD1 zymogen were identified by mass spectrometry. The autoactivation reaction with 0.5 μM zymogen (see above) was incubated for 2 h and stopped with 1 μM pepstatin. Peptide products were analyzed by MALDI-TOF (UltrafleXtreme, Bruker Daltonik) in the reflectron mode using α-cyano-4-hydroxycinnamic acid as a matrix. Protein products were desalted on-line and analyzed by LC/ESI-QTOF (I-class and Synapt G2, Waters); the resulting raw spectra were deconvoluted by MaxEnt1 (Waters). N-terminal Edman sequencing was applied to the protein products after Laemmli SDS-PAGE separation and electroblotting of gels.

Synthesis and Processing of Propeptide Fragments

Propeptide-derived peptides were synthesized by Fmoc solid phase chemistry on an ABI 433A Peptide Synthesizer (Applied Biosystems) in the form of peptidyl amides with free or acetylated N-termini (Doleckova-Maresova et al., 2005; Horn et al., 2011). For processing analysis, 30 μM peptides were incubated with 100 nM mature IrCD1 at 26 °C for up to 16 h in 0.1 M sodium acetate

pH 4.0. The resulting fragments were purified by RP-HPLC over a 5 μm C18 column (Luna, Phenomenex) equilibrated in 0.1% TFA and eluted with a linear gradient of 90% acetonitrile in 0.1% TFA. The purified peptides were characterized by ESI mass spectrometry on a Q-Tof Micro (Waters) (Jilkova et al., 2014).

Microscale Thermophoresis (MST)

Fluorescently labeled peptide [1-8] was prepared by modification of the N-terminal amino group with NT-647-NHS fluorescent dye (NanoTemper Technologies). NT-647-NHS and [1-8] were incubated at 0.1 mM concentrations in DMF for 4 h at alkaline pH (maintained by addition of N,N-diisopropylethylamine), and the product was purified by RP-HPLC and characterized by mass spectrometry (see previous section). For MST analysis, a series of IrCD1 dilutions (0-670 nM) were incubated with labeled peptide [1-8] (20 nM) in 50 mM Tris-HCl pH 7.0, 0.1 M NaCl, 0.05% Tween 20 for 15 min at 26 °C. The reaction mixture was measured in standard treated capillaries (NanoTemper Technologies) on a Monolith NT.115 instrument (NanoTemper Technologies) using 50% LED power and 20% IR laser power. The K_d value was determined by nonlinear fitting of the MST response using NT Analysis software 1.5.41. For competition experiments, IrCD1 (10 nM) was preincubated for 15 min at 26 °C with labeled peptide [1-8] (10 nM) in 50 mM Tris-HCl pH 7.0, 0.1 M NaCl, 0.05% Tween 20. Unlabeled peptide [1-8] was added at various concentrations (0-5 μM), and MST was measured as above.

In Silico Modeling of Peptide-IrCD1 Complexes

Based on the structure of peptide [1-8] bound to IrCD1, starting templates with substituted peptides were generated using the Rosetta fixed-backbone design protocol (Kuhlman and Baker, 2000). These templates were optimized by minimizing all internal degrees of freedom of the peptide, the active site side chains, and the rigid body orientation of the peptide using the FlexPepDock protocol (Raveh et al., 2010), as implemented in the FlexPepBind approach (Alam and Schueler-Furman, 2017; Alam et al., 2016; London et al., 2011). The resulting models were ranked based on their reweighted scores, which upweigh the contribution of the peptide to the binding energy (Raveh et al., 2011).

Rosetta version 9dec2014 was used for all runs in this study, using the Talaris energy function (Leaver-Fay et al., 2013):

(1) Substitution of peptide sequences using the Rosetta fixed-backbone protocol:

$ROSETTA_BIN/fixbb.linuxgccrelease -database $ROSETTA_DB -s complex.pdb -resfile mut_resfile -ex1 -ex2aro –use_input_sc -nstruct 1

where the Rosetta executable ($ROSETTA_BIN/fixbb.linuxgccrelease) and the Rosetta database ($ROSETTA_DB) paths are provided. The input structure is provided using the '*-s complex.pdb*' flag; substitution instructions are defined in the resfile and provided using the '*–resfile mut_resfile*' flag. Increased rotamer sampling is enforced for chi1 and aromatic chi2 angles, using the '*–ex1 –ex2aro*' flags. One output structure is generated as instructed by the '*–nstruct 1*' flag.

Given below is an example resfile to substitute the leucine residue of ^{1}AFRIP**L**TR^{8} at position 6 with methionine (given the chain id is B):

```
NATRO
start
6 B PIKAA M EX 1 EX 2 USE_INPUT_SC
```

(2) Minimization of structures using Rosetta FlexPepDock protocol:

$ROSETTA_BIN/FlexPepDocking.linuxgccrelease -database $ROSETTA_DB -s mutated_complex.pdb -flexPepDockingMinimizeOnly –nstruct 1

Minimization of the input mutated_complex.pdb is performed using the '*-flexPepDockingMinimizeOnly*' flag.

QUANTIFICATION AND STATISTICAL ANALYSIS

Enzyme kinetics and inhibition data were analysed using Grafit software (Erithacus) or NT Analysis software 1.5.41 (NanoTemper Technologies) with built-in algorithms. Means and standard errors have been derived from the best fit of the data, or based on three independent measurements, unless noted otherwise. For statistical software and analysis details, see Method Details and figure legends.

DATA AND SOFTWARE AVAILABILITY

The full list of software is presented in the Key Resources Table and in Method Details. The atomic coordinates and experimental structure factors have been deposited in the Protein Data Bank with codes 5N71, 5N7N, 5N70, and 5N7Q for mature IrCD1, IrCD1 zymogen, IrCD1-[1-8] complex, and IrCD1-pepstatin complex, respectively.

Cell Chemical Biology, Volume *25*

Supplemental Information

Novel Structural Mechanism of Allosteric Regulation of Aspartic Peptidases via an Evolutionarily Conserved Exosite

Iva Hánová, Jiří Brynda, Radka Houštecká, Nawsad Alam, Daniel Sojka, Petr Kopáček, Lucie Marešová, Jiří Vondrášek, Martin Horn, Ora Schueler-Furman, and Michael Mareš

SUPPLEMENTAL FIGURES

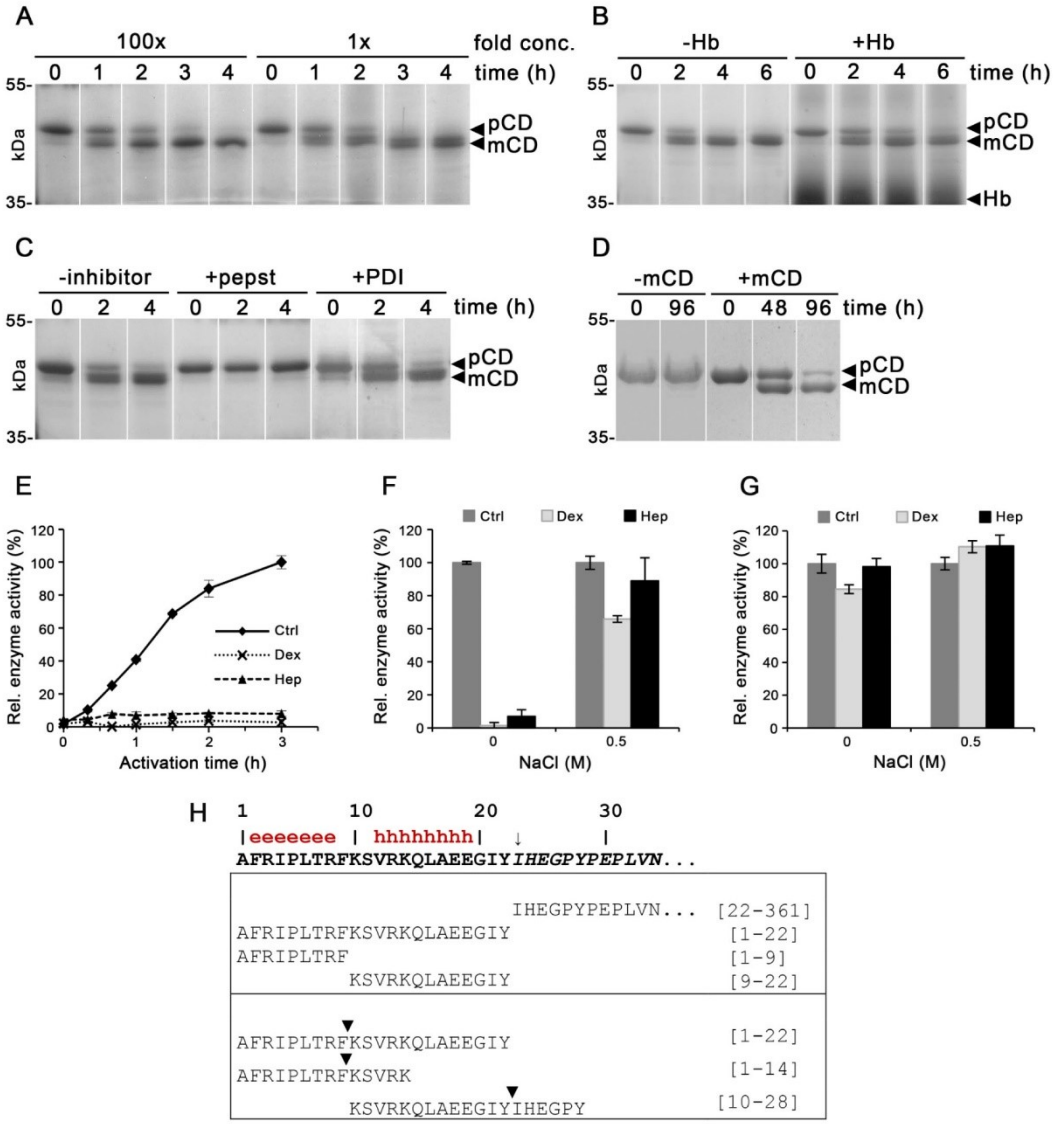


Figure S1. Processing of the IrCD1 Zymogen and its Propeptide, Related to Figure 1

(A-D) Molecularity of the autocatalytic processing of the IrCD1 zymogen. The IrCD1 zymogen was incubated at pH 4.0, and the reaction mixture was resolved by SDS-PAGE and visualized by protein staining. The positions of the zymogen (pCD) and mature enzyme (mCD) are marked.

(A) The effect of the zymogen concentration on the autoprocessing. The wild-type zymogen was incubated at 0.05 μM and 5 μM concentrations (1x and 100x, respectively) and analyzed at the indicated times.

(B) The effect of hemoglobin, a substrate of IrCD1, on the zymogen autoprocessing. The wild-type zymogen was incubated in the absence (-Hb) and presence (+Hb) of hemoglobin applied in a 3-fold

molar excess, and was analyzed at the indicated times. The position of hemoglobin oligomers is marked (Hb).

(C) The effect of active-site inhibitors of IrCD1 on the zymogen autoprocessing. The wild-type zymogen was incubated in the absence and presence of the peptidomimetic inhibitor pepstatin (+pepst) or 22-kDa potato cathepsin D inhibitor (+PDI) applied in a 10-fold molar excess and analyzed at the indicated times. Note that the autoprocessing is suppressed only by the low-MW inhibitor pepstatin interacting with the active site of cathepsins D in both mature and zymogen forms (Conner, Biochem J *263*, 601-604, 1989).

(D) The intermolecular processing of the active-site (D249N) mutated zymogen by exogenously added mature wild-type enzyme. The D249N zymogen was incubated in the absence (-mCD) and presence (+mCD) of purified active enzyme (200:1 molar ratio) for the indicated time. Note that the generated mCD band is derived from the D249N zymogen because exogenous mature enzyme was added in a catalytic amount.

(E-G) Autoactivation of the IrCD1 zymogen is suppressed by sulfated polysaccharides. Sulfated polysaccharides abolished autoactivation of the zymogen through ionic interactions (in a salt-sensitive manner), but they did not influence the activity of the mature enzyme.

(E) The zymogen was incubated at pH 4.0 in the absence (Ctrl) or presence of dextran sulfate (Dex) or heparin (Hep) (1 μg ml^{-1}). The activity of the mature enzyme was assayed at different time points in a kinetic assay at pH 4.0 with a FRET-peptide substrate. Mean values ± SE of triplicates are expressed relative to the maximum value.

(F) The zymogen was autoactivated for 1 h in the absence or presence of sulfated polysaccharides, and the mature enzyme was assayed as in (E). The experiment was performed in the absence or presence of 0.5 M NaCl to abolish ionic interactions with sulfated polysaccharides. Mean values ± SE of triplicates are expressed relative to the control without sulfated polysaccharides (100%).

(G) The mature enzyme was treated and analyzed as in (F) to demonstrate that its activity is not significantly affected by sulfated polysaccharides under the assay conditions.

(H) Cleavage map of the IrCD1 propeptide. The amino acid sequence of the IrCD1 propeptide is shown with secondary-structure elements derived from the crystallographic model (h, α-helix; e, β-strand). Residues corresponding to the N-terminal sequence of the mature enzyme are in italics; the autoprocessing site at the propeptide-enzyme core junction is indicated with an arrow. (Upper panel) Proteolytic products of the autoactivation reaction of the IrCD1 zymogen at pH 4.0 were determined by mass spectrometry, including the mature enzyme (residues 22-361), the full-length propeptide (residues 1-22), and two complementary propeptide fragments. (Lower panel) The cleavage pattern in synthetic peptides derived from the IrCD1 propeptide and the following mature N-terminus. The peptides were digested by mature IrCD1 at pH 4.0 (300:1 molar ratio), and the cleavage sites (black triangles) were identified by mass spectrometry. The peptides were synthesized with capping groups, N-terminal acetyl (for [10-28]) and C-terminal amide, to mimic peptide linkages.

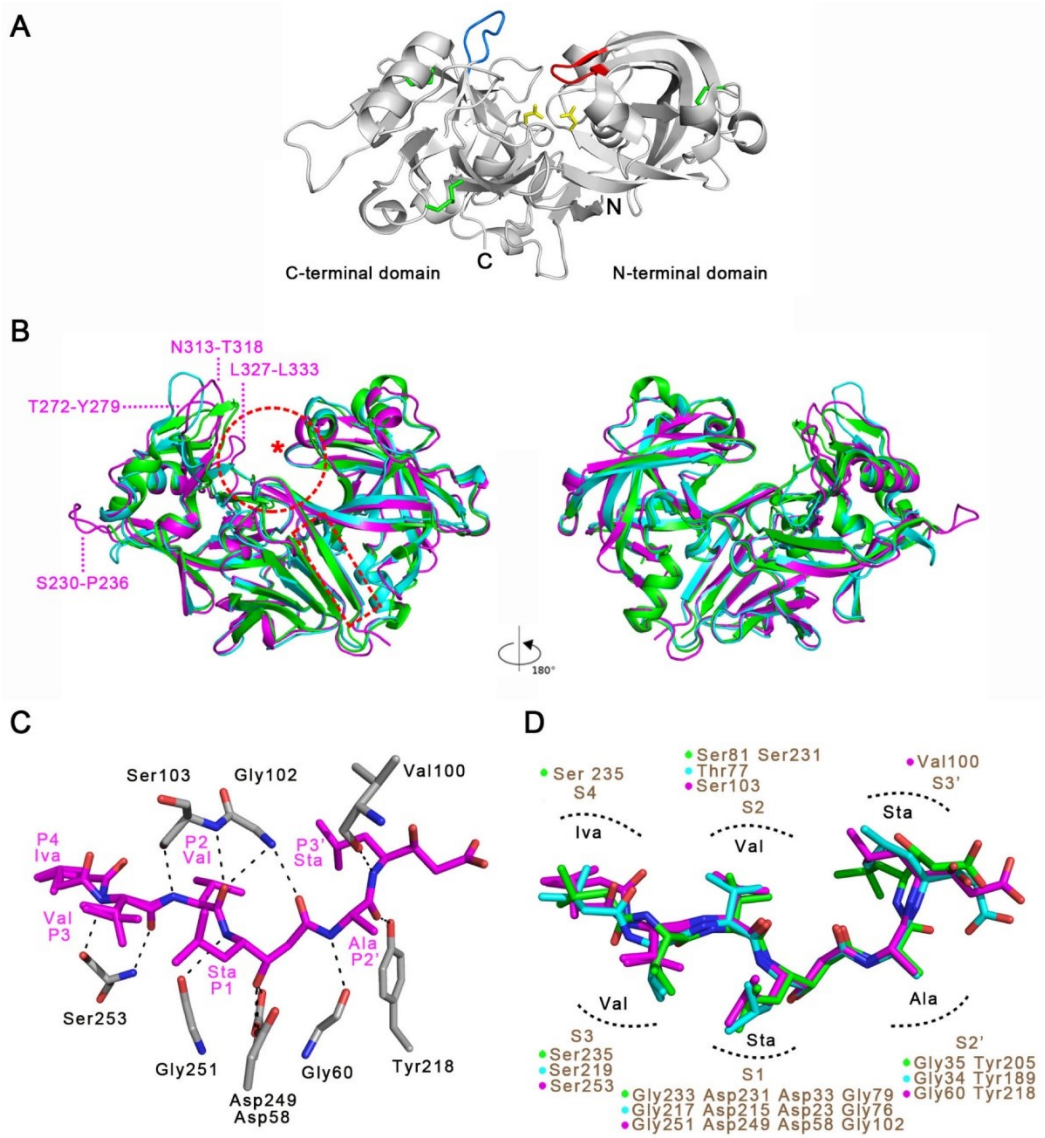


Figure S2. Crystal Structure of Mature IrCD1 and its Pepstatin Complex, Related to Figure 2

(A) The 3D structure of mature IrCD1 in the front view is in cartoon representation; the N- and C-termini and the N- and C-terminal domains are indicated. Important structural features of aspartic peptidases are highlighted: catalytic aspartates are in yellow sticks, the 'polyproline loop' is in dark blue, and the 'tyrosine flap' in red. Disulfide bridges are shown in green.

(B) Structural superposition of mature IrCD1 (magenta) with human cathepsin D (green) and bovine pepsin (cyan) (PDB codes 1LYA and 5PEP, respectively) is shown in cartoon representation. Locations of the active site (circled) and the exosite (boxed) are indicated in the front view. The 'tyrosine flap' is conserved in all structures and is marked with a red asterisk. Structurally divergent loops of IrCD1 are labeled with residue numbers.

(C) The substrate-analog inhibitor pepstatin (magenta) and the IrCD1 active-site residues (gray) forming hydrogen bonds (dashed lines) are in stick representation. Heteroatoms have standard color coding (N, blue; O, red). Pepstatin residues and their positions (Px to Px' nomenclature) in the active site are

labeled in magenta. The enzyme residues are labeled in black; Asp57 and Asp249 are catalytic residues. Non-canonical residues: Iva, isovaleryl; Sta, statine.
(D) Superposition of pepstatin molecules bound to the active sites of mature IrCD1 (magenta), human cathepsin D (green, PDB code 1LYA), and pepsin (blue, PDB code 1PSO) in stick representation (heteroatoms are colored as above). Pepstatin residues are labeled in black. Enzyme residues forming hydrogen bonds with pepstatin and their location in the subsites (dotted curves with Sx to Sx' nomenclature) are labeled in gray; color-coded markers indicate enzymes (magenta, IrCD1; green, human cathepsin D; blue, pepsin). Asp residues in the S1 subsite are catalytic residues.

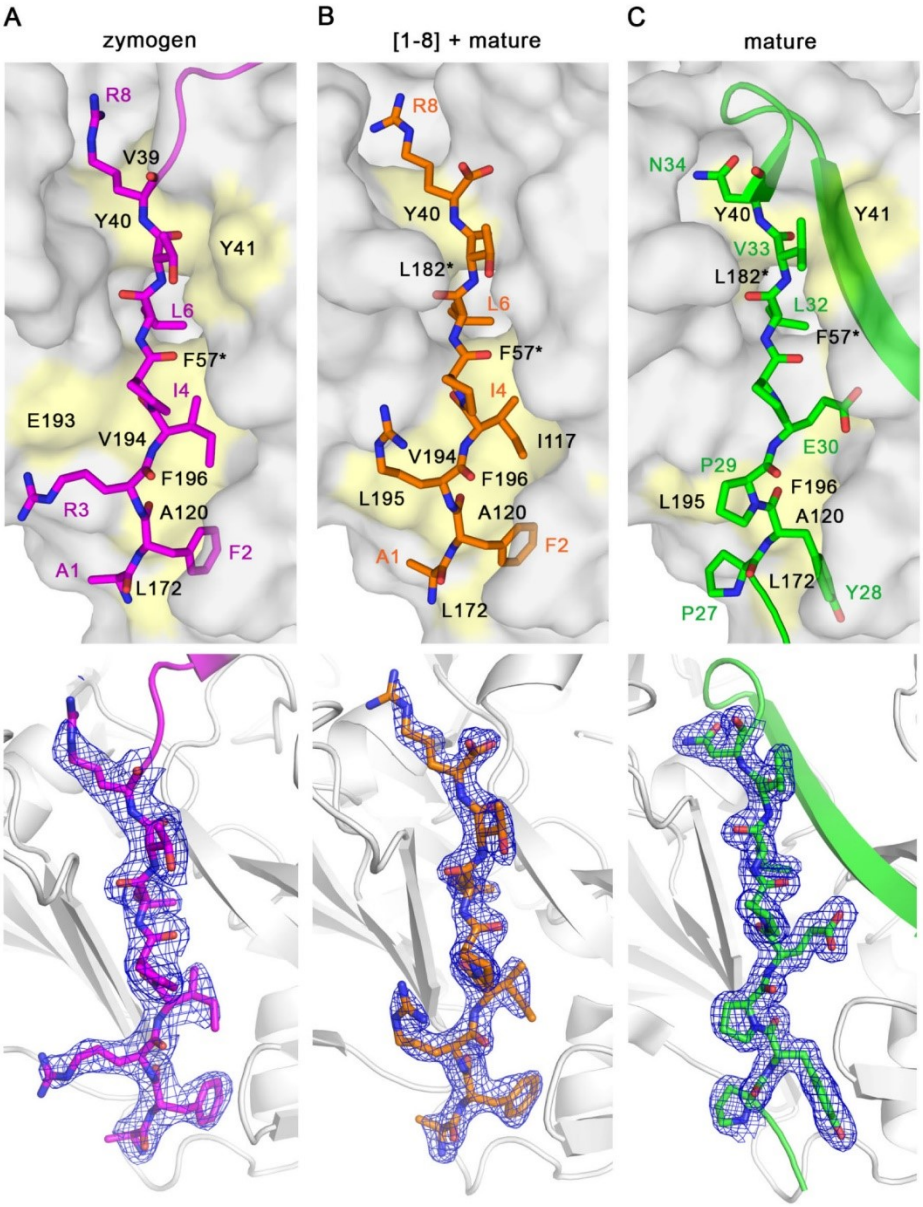


Figure S3. Comparison of IrCD1 Exosite Ligands, Related to Figure 4

(Upper panels) Nonpolar interactions of three different ligands in the same exosite are shown for: (a) the N-terminal region (residues 1-8) (magenta) of the intact propeptide in the structure of the IrCD1 zymogen, (b) the synthetic peptide derived from propeptide residues 1-8 (orange) in complex with mature IrCD1, and (c) the mature N-terminus (residues 26-34) (green) in the structure of mature IrCD1. The ligands are in stick representation (heteroatoms have standard color coding: O, red; N, blue) followed by a stretch in cartoon representation (colored as for ligands). The enzyme core is shown as a gray surface with the residues forming nonpolar interactions with ligands in yellow. Ligand residues forming nonpolar interactions are labeled in color; the enzyme core residues are labeled in black. Residues labeled with an asterisk are hidden inside the cavity.

(Lower panels) Electron density map of IrCD1 exosite ligands. The $2F_o$-F_c electron density maps contoured at 1.2 σ are shown. The ligands are in stick representation followed by a stretch in cartoon representation (colored as for ligands); the enzyme core is shown as a gray cartoon.

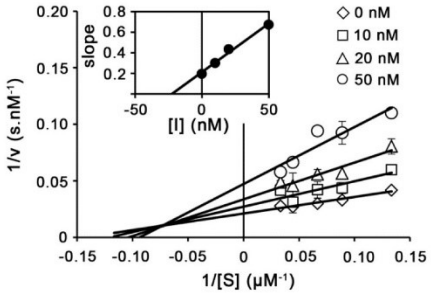


Figure S4. Kinetic Analysis of IrCD1 Inhibition by Synthetic Propeptide, Related to Figure 6

The Lineweaver-Burk plot for inhibition of mature IrCD1 with the full-length IrCD1 propeptide is indicative of a mixed-type, noncompetitive inhibition (the family of straight lines intersects in the second quadrant). The inset shows a secondary plot of the same data; an inhibition constant of K_i = 22 nM was determined. Enzyme activity was measured in a kinetic assay with a FRET-peptide substrate (the IrCD1-inhibitor complex was preformed at pH 7 and assayed at pH 5). Error bars depict the mean ± SE from triplicate measurements.

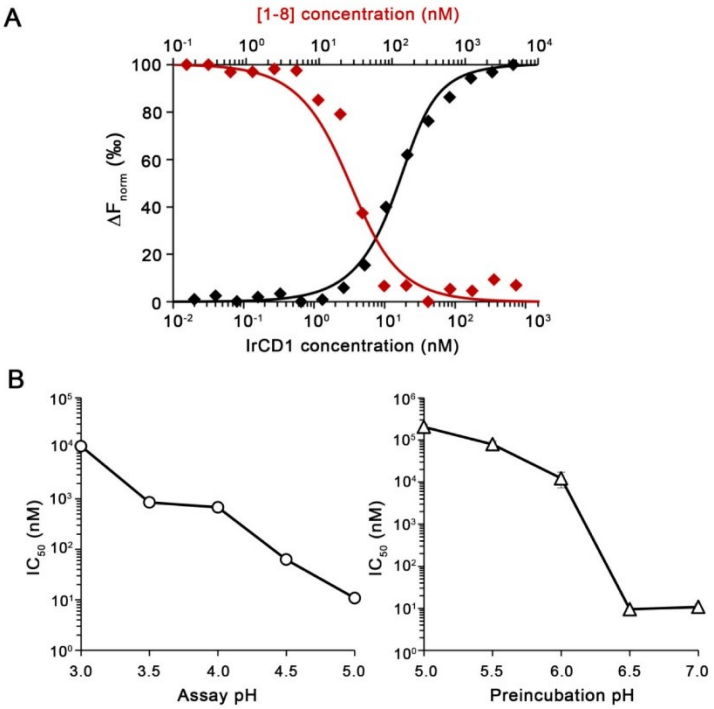


Figure S5. Interaction of the Propeptide Fragment [1-8] with IrCD1, Related to Figure 6

(A) Microscale thermophoresis analysis. Titration of mature IrCD1 against a constant concentration of fluorescently labeled fragment [1-8] (black curve) revealed an optimized K_d of 6.8 ± 0.9 nM. Competitive displacement of the fluorescently labeled fragment [1-8] complexed with mature IrCD1 with unlabeled fragment [1-8] (red curve) revealed an optimized K_d of 3.4 ± 1.0 nM. The concentration of titrant is plotted against the normalized value of fluorescence change between the unbound and bound states. Measurements were performed at pH 7.0 in duplicates; a representative result is shown. K_d values were calculated using NanoTemper Analysis software.

(B) pH-dependent inhibition of IrCD1 by [1-8]. (Left panel) Mature IrCD1 was preincubated with [1-8] at pH 7 for 10 min, and enzyme activity was then measured in a kinetic assay with a FRET-peptide substrate at pH 3-5. The IC_{50} values were determined; mean values ± SE are given for triplicate measurements. (Right panel) Mature IrCD1 was preincubated with [1-8] at pH 5-7 for 10 min, and enzyme activity was then measured in a kinetic assay with a FRET-peptide substrate at pH 5. The IC_{50} values were determined; mean values ± SE are given for triplicate measurements.

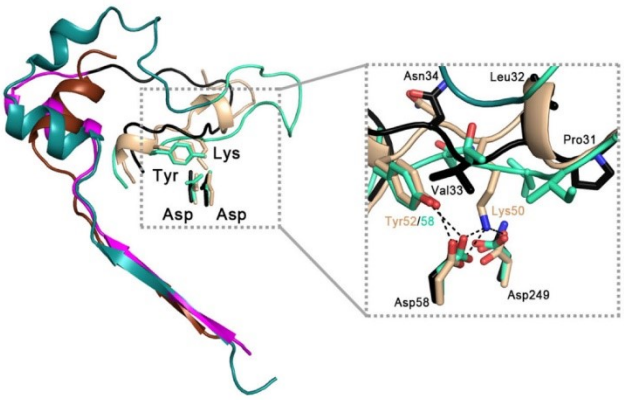


Figure S6. Structural Homology of the IrCD1 Propeptide, Related to Figure 3

Structural superposition of the propeptides of the IrCD1 zymogen, and human prorenin and plant prophytepsin (PDB codes 3VCM and 1QDM, respectively) in cartoon representation; for sequence alignment, see Figure 5. The propeptide and the following mature region are shown in magenta and black, respectively, for the IrCD1 zymogen, dark green and mint for prorenin, and brown and beige for prophytepsin (part of which is not resolved). The zoom view shows the interaction of the mature regions of prorenin and prophytepsin (through the Lys and Tyr residues) with catalytic aspartate residues of the active site. The interacting residues are depicted as sticks and direct hydrogen bonds as dashed lines. Note that this interaction mechanism is absent in IrCD1 (black).

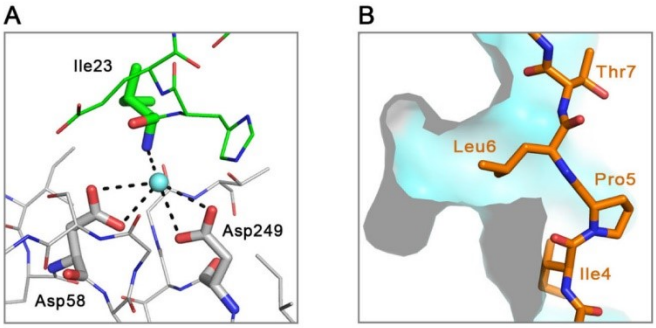


Figure S7. Detail of Critical Inhibitory Interactions in the Complex of Mature IrCD1 with the Inhibitor [1-8], Related to Figures 3 and 4

(A) View of the enzyme active site. The N-terminal residue Ile23 (green sticks) of the mature enzyme forms an indirect hydrogen bond (dashed lines) with catalytic residues Asp58 and Asp249 (grey sticks) through a water molecule (blue sphere).

(B) View of the enzyme exosite. The Leu6 residue of the [1-8] inhibitor (orange sticks) is critical for its interaction with the exosite (pale blue surface). Leu6 is inserted into a deep hydrophobic pocket, which is shown in a cut-open view. Heteroatoms have standard color coding (N, blue; O, red).

SUPPLEMENTAL TABLES

Table S1. Data Collection and Refinement Statistics, Related to Figure 2

	mature IrCD1	mature IrCD1 + pepstatin	IrCD1 zymogen	mature IrCD1 + [1-8]
PDB code	**5N71**	**5N7Q**	**5N7N**	**5N70**
Data collection				
Space group	$P4_12_12$	$P2_1$	$P2_1$	$P4_1$
Cell dimensions				
a, *b*, *c* (Å)	98.16, 98.16, 90.80	61.14, 82.82, 61.39	95.32, 94.98, 106.17	63.92, 63.92, 80.50
α, β, γ (°)	90, 90, 90	90, 91.1, 90	90, 99.3, 90	90, 90, 90
Resolution (Å)	49.08-1.88 (1.99-1.88)	61.38-1.46 (1.53-1.46)	45.87-2.30 (2.44-2.30)	45.20-1.80 (1.90-1.80)
*R*merge	0.09 (0.95)	0.08 (0.23)	0.07 (0.48)	0.05 (0.63)
$I / \sigma I$	16.16 (2.33)	7.20 (3.10)	14.81(2.79)	16.18 (2.43)
Completeness (%)	99.6 (98.0)	98.0 (98.6)	99.7 (99.3)	99.4 (96.5)
Redundancy	8.0 (7.8)	3.0 (2.8)	3.9 (3.9)	6.0 (5.9)
Refinement				
Resolution (Å)	49.08-1.88 (1.93-1.88)	61.38-1.46 (1.49-1.46)	45.87-2.30 (2.36-2.30)	45.20-1.81 (1.85-1.81)
No. reflections	36,744 (2,228)	101,593 (7,296)	29,567 (1,845)	82,782 (5,583)
*R*work / *R*free	0.198 (0.224)/ 0.236 (0.281)	0.146 (0.205)/ 0.166 (0.229)	0.220 (0.296)/ 0.260 (0.347)	0.182 (0.247)/ 0.228 (0.301)
No. atoms				
Protein	2,568	5,214	10,968	2,659
Ligand/ion	12	111	94	0
Water	122	1309	147	94
No. molecules in AU	1	2	4	1
B-factors				
Protein	31.0	13.6	38.3	41.5
Ligand/ion	53.0	20.6	72.4	-
Water	29.9	28.3	26.6	37.2
R.m.s deviations				
Bond lengths (Å)	0.014	0.010	0.009	0.017
Bond angles (°)	1.55	1.50	1.45	1.9

For each structure, a single crystal was used.
Values in parentheses are for the highest resolution shell.

Table S2. List of Contacts Formed between the IrCD1 Active Site and Pepstatin, Related to Figure 3

Position	Residue	Contacts	H-bonds
P4	Val 38	2	
	Ser 103	1	
	Ser 253	3	
	Leu 254	2	
	Gln 278	2	
	Leu 310	2	
P3	Val 39	2	
	Ser 103	7	
	Gly 251	4	
	Thr 252	5	
	Ser 253	10	1,1
P2	Tyr 101	1	
	Gly 102	3	1
	Ser 103	7	1,1
	Gly 251	1	
	Thr 252	3	
	Leu 327	1	
P1	Ile 56	2	
	Asp 58	10	1
	Gly 60	3	
	Tyr 101	7	
	Gly 102	1	1
	Asp 249	5	2
	Gly 251	9	1
	Thr 252	2	
P2'	Ser 61	1	
	Gly 60	4	1
	Val 100	3	
	Tyr 101	2	
	Tyr 218	4	1
P3'	**Val 100**	5	1
	Gly 102	3	
	Tyr 218	2	
	Met 329	2	

Analysis of contacts was performed using the program CCP4. The distance cutoffs were set to 3.3 Å and 4.0 Å for hydrogen bonds (H-bonds) and van der Waals interactions, respectively. The IrCD1 residues interacting with individual pepstatin residues (positions P3 to P3') are specified. For each IrCD1 residue, the total number of contacts, including H-bonds, is indicated. The H-bond numbers are colored according to the type of IrCD1-pepstatin interaction (red: backbone-backbone, green: side chain-backbone).

Table S3. List of Contacts Formed between the IrCD1 Exosite and its Ligands, Related to Figure 4

	[1-8]			N-terminus zymogen			N-terminus mature		
Position	**Residue**	**Contacts**	**H-bonds**	**Residue**	**Contacts**	**H-bonds**	**Residue**	**Contacts**	**H-bonds**
P1^E	Leu 195	1					Leu 195	1	
	Phe 196	3		Phe 196	2		Phe 196	3	
	Gly 197	2					Gly 197	4	
							Gly 198	2	
P2^E	Ile 117	2		Ile 117	1				
	Ser 119	1		Ser 119	1		Ser 119	2	
	Ala 120	4		Ala 120	6		Ala 120	10	
	Met 166	2		Met 166	1				
	Leu 171	1		Leu 171	1		**Leu 171**	5	1
	Leu 172	2		Leu 172	1		Leu 172	1	
							Val 193	1	
	Val 194	1							
	Leu 195	4		Leu 195	3		Leu 195	4	
	Phe 196	14	2	**Phe 196**	11	2	**Phe 196**	9	2
P3^E	**Glu 193**	9	2	**Glu 193**	9	1	Glu 193	2	
	Val 194	2		Val 194	2		Val 194	2	
							Leu 195	3	
				Asp 200	3	1			
P4^E							Gly 42	3	
	Pro 43	1		Pro 43	1		**Pro 43**	8	3
	Phe 57	2		Phe 57	2		Phe 57	6	
	Ile 117	3		Ile 117	1		Ile 117	1	
	Gly 118	2		Gly 118	2				
	Glu 193	3		Glu 193	4		Glu 193	4	
	Val 194	4	1	**Val 194**	4	2	**Val 194**	8	2
	Phe 196	1							
P5^E	Gly 42	1		Gly 42	1		Gly 42	1	
	Pro 43	1		Pro 43	1		Pro 43	1	
	Gly 192	2		Gly 192	2		Gly 192	2	
P6^E	Tyr 41	5		Tyr 41	5		Tyr 41	5	
				Gly 42	1		Gly 42	1	
	Phe 57	1		Phe 57	1		Phe 57	1	
	Leu 182	1					Leu 182	1	
	Gly 191	4	1	**Gly 191**	3	1	**Gly 191**	3	1
	Gly 192	7	1	**Gly 192**	11	2	**Gly 192**	9	1
	Glu 193	9		Glu 193	5		Glu 193	2	
	Val 194	2		Val 194	1				
P7^E	Val 39	1		Val 39	1		Val 39	1	
	Tyr 40	2		Tyr 40	2		Tyr 40	4	
	Tyr 41	13	2	**Tyr 41**	9	2	**Tyr 41**	10	2
P8^E				Phe 9	4				
							Leu 36	4	
	Asp 37	1	1				**Asp 37**	6	1,1
	Val 38	6							
	Val 39	4	1	Val 39	4		Val 39	2	
	Tyr 40	9		Tyr 40	9		Tyr 40	10	
	Tyr 41	1							
				Asn 183	1		**Asn 183**	2	1
	Arg 184	4	1	Arg 184	3				
							Pro 186	1	
	Gly 191	1					Gly 191	2	

The exosite ligands are the propeptide-derived octapeptide inhibitor [1-8], the N-terminal propeptide segment (residues 1-8) of the zymogen, and the N-terminal segment (residues 26-34) of the mature enzyme. Analysis of contacts between the ligands and the enzyme core was performed using the program CCP4. The distance cutoffs were set to 3.3 Å and 4.0 Å for hydrogen bonds (H-bonds) and van der Waals interactions, respectively. The IrCD1 residues interacting with the individual residues (positions P1^E to P8^E) of the exosite ligands are specified. For each IrCD1 residue, the total number of contacts, including H-bonds, is indicated. The H-bond numbers are colored according to the type of IrCD1-ligand interaction (red: backbone-backbone, yellow: backbone-side chain, blue: side chain-side chain).

Table S4. Predicted Binding Energies for Exosite Inhibitors Substituted at the Leu6 Position, Related to Figure 6

P6 Amino acid	Δ Reweighted score (REU)
L	**0.0**
M	5.0
Q	6.7
I	8.9
C	10.2
V	10.2
E	10.8
N	11.0
T	11.5
A	11.6
F	11.9
H	12.7
D	13.6
S	14.3
R	14.6
G	15.4
K	15.5
Y	18.9
W	19.1
P	45.9

Derivatives of the inhibitor [1-8] incorporating systematic substitutions with canonical amino acids at the Leu6 position were modeled in complex with IrCD1. Binding energies were calculated using the Rosetta FlexPepDock protocol (see STAR Methods) and expressed in Rosetta Energy Units (REU). Reweighted scores are given relative to the wild-type inhibitor (red) (Δ Reweighted score).

4.3 Publikace č. 3: Biomimetic macrocyclic inhibitors of human cathepsin D: structure-activity relationship and binding mode analysis

4.3.1 Souhrn

Proti KatD jako cílové molekule pro terapii jsou vyvíjeny *de novo* syntetické inhibiční regulátory[212]. Tato studie vychází z biomimetického přístupu a navrhuje nové inhibitory KatD na bázi mikrobiálního peptidu pepstatinu. Ten byl původně izolován z bakterií rodu *Actinomyces* a je nejúčinnějším přirozeným inhibitorem aspartátových proteas. Lineární molekula pepstatinu obsahuje aminokyselinu Sta, která interaguje s katalytickými aspartáty. Nicméně pepstatin, podobně jako další přirozené lineární peptidy, není vhodný pro medicinální využití vzhledem k obecně nepříznivým fyzikálně-chemickým a farmakokinetickým vlastnostem[213].

Na prostorové struktuře komplexu KatD s pepstatinem je patrné, že se pepstatin v aktivním místě enzymů stáčí a zaujímá tvar „U“ s aminokyselinovými zbytky P2 a P3` v ramenech blízko u sebe. S cílem stabilizovat tuto vazebnou konformaci byly pozice ramen propojeny alifatickou spojkou za vzniku makrocyklické struktury. Na jejím základě byla připravena sada více než 30 modifikovaných derivátů s výrazně menší molekulovou hmotností oproti původnímu pepstatinu. Optimalizována byla struktura spojky a testovány substituce v několika polohách včetně nových derivátů Sta.

Získány byly silné reverzibilní inhibitory KatD s až subnanomolárními inhibičními parametry a s kompetitivním módem inhibice. Selektivita inhibitorů byla testována se sedmi zástupci třídy aspartátových proteas. Inhibitory vykazovaly obecně výraznou selektivitu pro strukturně blízké proteasy typu KatD nebo pepsinu z rodiny A1 oproti evolučně vzdálenějším proteasám rodiny A1 a A2 (jako jsou renin, BACE-1 nebo HIV proteasa). Zároveň byla prokázána možnost účinně ladit inhibiční specifitu pomocí substitucí.

Pro tři inhibitory byly připraveny komplexy s lidským KatD a byly řešeny jejich prostorové struktury rentgenostrukturní analýzou. Ukázaly, že makrocyklické inhibitory fungují jako konformační mimetika vazebného módu pepstatinu, přičemž jejich minimalizovaná stavba vytváří analogický počet hlavních kontaktů s aktivním centrem jako větší molekula pepstatinu. S využitím prostorových struktur byly metodami výpočetní

chemie analyzovány interakční energie a příspěvky jednotlivých částí makrocyklů k inhibici.

Vyvinuté makrocyklické inhibitory splňují kritéria fyzikálně-chemických parametrů pro navrhování léčiv, jsou proteolyticky stabilní a netoxické a představují perspektivní templát pro vývoj chemoterapeutik.

Můj podíl na práci zahrnoval: (1) měření inhibic vybraných makrocyklických inhibitorů s modelovými proteasami, (2) analýzu proteolytické stability a fyzikálně-chemických parametrů inhibitorů, (3) přípravu manuskriptu.

4.3.2 Publikace č. 3

Journal of Medicinal Chemistry

pubs.acs.org/jmc Article

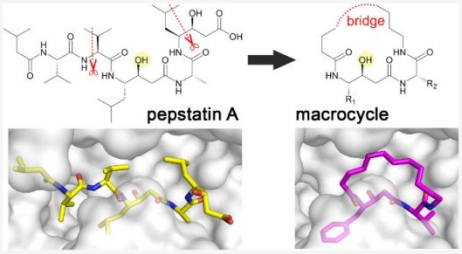
Biomimetic Macrocyclic Inhibitors of Human Cathepsin D: Structure–Activity Relationship and Binding Mode Analysis

Radka Houštecká, Martin Hadzima, Jindřich Fanfrlík, Jiří Brynda, Lenka Pallová, Iva Hánová, Helena Mertlíková-Kaiserová, Martin Lepšík, Martin Horn, Martin Smrčina, Pavel Majer,* and Michael Mareš*

Cite This: J. Med. Chem. 2020, 63, 1576−1596 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Human cathepsin D (CatD), a pepsin-family aspartic protease, plays an important role in tumor progression and metastasis. Here, we report the development of biomimetic inhibitors of CatD as novel tools for regulation of this therapeutic target. We designed a macrocyclic scaffold to mimic the spatial conformation of the minimal pseudo-dipeptide binding motif of pepstatin A, a microbial oligopeptide inhibitor, in the CatD active site. A library of more than 30 macrocyclic peptidomimetic inhibitors was employed for scaffold optimization, mapping of subsite interactions, and profiling of inhibitor selectivity. Furthermore, we solved high-resolution crystal structures of three macrocyclic inhibitors with low nanomolar or subnanomolar potency in complex with CatD and determined their binding mode using quantum chemical calculations. The study provides a new structural template and functional profile that can be exploited for design of potential chemotherapeutics that specifically inhibit CatD and related aspartic proteases.



INTRODUCTION

Cathepsin D (CatD), a ubiquitous aspartic protease, is involved in several important physiological and pathological processes in humans. In addition to its primary role in lysosomal digestion of proteins, CatD performs a variety of other specialized functions, including processing of peptide hormones, activation of enzymatic precursors, antigen presentation, degradation of cytoskeletal proteins, matrix remodeling, and regulation of apoptosis.[1] CatD is implicated in numerous pathologies, including several neurodegenerative conditions, osteoarthritis, and acute pancreatitis.[1−3] However, it has mainly been studied in connection with different types of carcinomas. CatD is overexpressed and hypersecreted by breast cancer cells and is associated with tumor development and progression.[4,5] It serves as an independent marker of poor breast cancer prognosis and is correlated with the incidence of clinical metastasis.[6,7] Therefore, CatD represents a promising target for the treatment of breast cancer using antibody-based therapy or chemotherapy.[8,9]

Various aspartic proteases are targets for drugs that regulate their proteolytic activity. Renin inhibitors are used to treat hypertension and HIV protease inhibitors to treat HIV infection, whereas β-secretase (BACE1) and plasmepsins are being investigated as targets for drugs against Alzheimer's disease and malaria, respectively.[10−13] Several types of potential small-molecule CatD inhibitors have been designed de novo.[14−17] The most potent natural inhibitor of CatD discovered to date is pepstatin A.[18] This linear pseudo-pentapeptide originally isolated from actinobacteria is effective against many aspartic proteases from the pepsin family with inhibition constants as low as 10^{-10} M. Its potency is attributable to the nonproteinogenic amino acid statine [(3*S*,4*S*)-4-amino-3-hydroxy-6-methylheptanoic acid], the hydroxyl group of which interacts with catalytic aspartate residues and acts as a transition-state analogue. Pepstatin A is widely used as a research tool for enzyme affinity purification, active site titration, imaging, and structural mapping of binding sites.[19−21] However, its use as a therapeutic is limited due to its large size and unfavorable physicochemical properties including poor solubility, which are reflected in its low bioavailability and rapid clearance.[22]

In this work, we used the spatial conformation that pepstatin A adopts upon binding to the CatD active site to design a macrocyclic mimetic with minimized size and potentiated inhibition. This biomimetic scaffold was subjected to comprehensive structural characterization, including binding

Received: August 30, 2019
Published: January 31, 2020

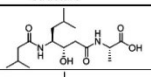
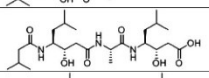
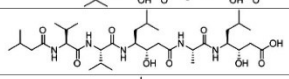
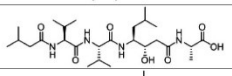
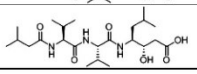
https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

mode analysis based on crystallography of CatD−inhibitor complexes and quantum chemical calculations. Furthermore, we performed a detailed structure−activity relationship (SAR) analysis of the scaffold using an extensive library of substituted macrocycles, leading to identification of CatD inhibitors with low nanomolar and subnanomolar potency. In summary, our study yielded a novel class of inhibitors inspired by natural compound but with improved drug-like properties. These new compounds represent a starting point for rational development of potential chemotherapeutics against pathologies associated with human CatD and CatD-like proteases in general.

RESULTS

Design of Macrocyclic Mimetics of the Natural Inhibitor Pepstatin A and SAR Analysis of Human CatD inhibition. *Truncation of the Linear Scaffold of Pepstatin A Decreases Inhibitory Potency.* A series of linear truncated fragments of pepstatin A (compound **1**) was designed to identify portions of the inhibitor that can be removed without loss of potency toward human CatD (Table 1). Removal of the C-terminal P3′ statine from pepstatin A (**2**)

Table 1. Effect of Pepstatin A (1) Truncation on Inhibitory Potency toward Human CatD

Compound	Structure	IC_{50} (nM)
5		>200,000
4		2,500 ± 440
1		0.49 ± 0.03
2		40 ± 5
3		500 ± 100

(Figure 1) led to a 2-order-of-magnitude drop in potency (IC_{50} of 0.5 vs 40 nM), and further trimming of the Ala residue resulted in an additional 1-order-of-magnitude decrease in potency (**3**). Truncation of the N-terminal isovaleryl-Val segment in the P4−P3 positions (**4**) resulted in a dramatic decrease in inhibition to the μM level. Compound **5**, which retained only the P2 to P2′ residues, showed very weak sub-mM inhibition. These observations indicate that a full-sized natural inhibitor spanning P4 to P3′ is needed to inhibit CatD in the subnanomolar range. Stepwise removal of substituents from both termini around the linear central segment with the P1 statine (Figure 1) leads to substantial reductions in inhibitory potency. This is in line with previous SAR data for bovine CatD.[23] In the next step, we explored the effect of cyclization on the truncated central segment.

Scaffold Cyclization Boosts Inhibition and Yields Macrocycles with Low and Subnanomolar Potency. When bound to human CatD, the pepstatin A molecule forms a U-shape with the P2 and P3′ side chains in proximity to each other (Figure 1). These side chains form contiguous interactions in

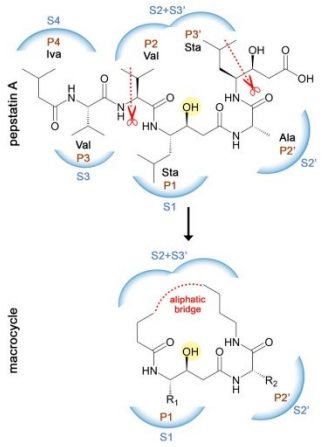


Figure 1. Schematic representation of the rational design of the macrocyclic inhibitor scaffold. The scaffold mimics the spatial conformation of pepstatin A, a natural linear oligopeptide inhibitor, in the active site of human CatD (PDB: 1LYB). Binding subsites (S) are marked and colored in blue; corresponding inhibitor positions (P) are in brown. The central hydroxyl group of the statine residue that interacts with the catalytic aspartates of CatD is highlighted in yellow. Lines with scissors indicate the region used for macrocycle scaffold design. Pepstatin A structure: Iva-Val-Val-Sta-Ala-Sta, where Iva is isovaleryl and Sta is statine.

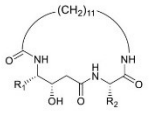
the corresponding S2 and S3′ subsites of the enzyme. S2 and S3′ are located in a markedly hydrophobic, spacious area that also includes S1′, the subsite for the P1′ side chain, which is missing in statine-based inhibitors. The absence of a P1′ substituent might be the reason that the P2 isopropyl and P3′ isobutyl bend toward each other to fill the empty space. The proximity of those substituents led us to the idea of connecting them covalently to form a macrocycle (Figure 1). We designed a P2−P3′ cycle to span a sufficient portion of the binding cavity and include the backbone of the P1 statine residue as a transition-state isostere. Most of the interaction free energy of pepstatin A with CatD comes from hydrophobic interactions, as demonstrated by the dominant contribution of the entropic term.[24] Therefore, we assumed that improving the hydrophobic interactions of the P2/P3′ substituents, along with organization into the presumed bound conformation, may compensate for the lost interactions caused by truncation (removal of residues outside of P2 and P3′) (Figure 1).

Our cyclic pseudo-tripeptide scaffold is formed by the P1 statine residue and the neighboring P2′ residue bridged by 12-aminododecanoic acid, which serves as a mimetic of the linked P2 and P3′ residues (Table 2a). The size of the bridge was based on molecular modeling, which indicated sufficient space in the active site to accommodate it. Compound **6** with the original pepstatin A substituents, P1 isobutyl and P2′ methyl, exhibited only ∼300 nM potency. The crystal structure of bound pepstatin A revealed that both these side chains only partially fill the subsites. We therefore introduced the bulkier benzyl moiety into P1 (**8**) and isopropyl into P2′ (**7**); these compounds were roughly an order of magnitude more potent than **6**. Simultaneous replacement in **9** led to further

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

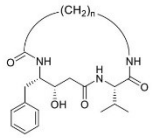
Table 2. Initial Design of Pepstatin A-Mimicking Macrocyclic Inhibitors of Human CatD[a]

a



Compound	R_1	R_2	IC_{50} (nM)
6			310 ± 61
7			22 ± 5
8			68 ± 16
9			5.8 ± 0.7

b



Compound	n	Ring size	IC_{50} (nM)
10	10	20	40 ± 9
9	11	21	5.8 ± 0.7
11	12	22	3.5 ± 0.5
12	13	23	11 ± 1
13	14	24	32 ± 4

[a]Compound **6** contains the original pepstatin A substituents at R_1 and R_2. Modification of these two positions (a) and optimization of macrocycle size (b) improved the inhibitory potency of the macrocyclic scaffold.

improvement in potency (IC_{50} of ~6 nM). Next, we optimized the size of the macrocycle and prepared a series of inhibitors with 10−14 methylene units (**9** to **13**). The IC_{50} values varied within 1 order of magnitude, with optimal inhibition by compounds with 11 or 12 methylene units (Table 2b). We chose the smaller ring size for further optimization of the P1 and P2′ substituents.

We introduced structural changes at the P1 and P2′ positions (i.e., substituents R_1 and R_2, respectively) to map the CatD subsite preferences, using **9** as a standard to which the IC_{50} values were compared. In the first series, we optimized the R_1 substituent by evaluating different statine analogues with a fixed valine (R_2 isopropyl) in the P2′ position (Table 3a). In general, hydrophobic aromatic residues were preferred in the P1 position. Compound **14** with *β*-naphthyl exhibited the lowest IC_{50} of 0.6 nM. Inhibitors with a single aromatic ring (**15**, **16**, **17**, and **9**) had slightly higher IC_{50} values. The minimal difference between **9** with benzyl and **18** with cyclohexyl-methyl suggested that hydrophobic interaction prevails over aromatic stacking with the several aromatic residues in the S1 pocket. Introduction of more hydrophilic residues (**20**, **22**, **23**, and **24**) led to substantial loss of potency. Hence, **22** with *β*-indolyl was 2 orders of magnitude weaker than **14** with *β*-naphthyl, and **24** with a smaller thiazolylmethyl moiety was the weakest compound of the entire series, 2 orders of magnitude weaker than **9** with benzyl.

The second series of compounds included a fixed phenylstatine residue in P1 (R_1 benzyl) and modifications at the P2′ position (Table 3b). The most potent compound **25** with 3-pyridyl alanine had an IC_{50} of approximately 1 nM, five times more potent than **9** with valine. Comparison of **25** and **28** with benzyl (Phe) in P2′ yielded an interesting result: it is the only example in the series in which introduction of a more polar residue increased the potency. In general, medium-size hydrophobic side chains were preferred (**26**, **9**, **27**, and **30**). Inhibitor **31** with small hydrophilic threonine was 2 orders of magnitude weaker than **9**, as was **32**, likely due to excessive steric volume in the relatively small P2′ pocket.

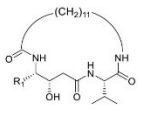
We also assessed various modifications along the macrocycle backbone (Table 4). The most active compound in this series—**33** with a urethane moiety—has the same cycle size (22 ring members) as the "full carbon" analogue **11** and exhibited similar potency. Replacement of amide with urea in **34** (also a 22-membered ring) led to a slight reduction in potency, whereas replacement of the polymethylene chain with an ethylene glycol-based linker led to a weak inhibitor (**35**) with an IC_{50} of 850 nM. This further highlights the importance of hydrophobic interactions in the P2/P3′ region. Compound **36**, which lacks a hydroxyl on the phenylstatine backbone, has a three-order-of-magnitude lower potency than its closest analogue **9**, which directly demonstrates the essential role of the hydroxyl group interacting with catalytic aspartates.

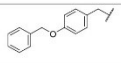
Macrocycles Have Drug-like Properties. We evaluated the basic physicochemical properties of representative macrocyclic inhibitors and calculated parameters of Lipinski's "rule of five" for drug-likeness to compare a set of the most potent macrocycles with pepstatin A (Supporting Information Table 1). Numbers of hydrogen bond donors and acceptors for macrocycles (4 and 7/8, respectively) meet the "rule of five" criteria (values lower than 5 and 10), while values for pepstatin A are roughly double that of macrocycles and do not meet the criteria. The molecular mass of macrocycles was in the range of 488−557 Da; the favorable values are below 500 Da; however, this parameter has previously been shown to be rather relaxed for macrocycle drugs.[25] Pepstatin A, with molecular mass of 686 Da, violated the rule. The partition coefficients for all compounds complied with the rule (log *P* below 5). Additional parameters used in drug design, including the number of rotatable bonds and topological polar surface area, were within the favorable range (below 10 and below 140 Å^2, respectively) for macrocycles (3 and 108−120 Å^2) and out of range for pepstatin A (22 and 233 Å^2) (Supporting Information Table 1).

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

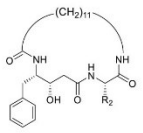
Table 3. Substitution Analysis of Macrocyclic Inhibitors of Human CatD[a]

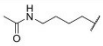
a



Compound	R_1	IC_{50} (nM)
14		0.6 ± 0.1
15		3.1 ± 0.6
16		4.2 ± 0.7
17		4.7 ± 0.7
9		5.8 ± 0.7
18		9.0 ± 0.7
19		20 ± 4
7		22 ± 5
20		39 ± 9
21		55 ± 5
22		190 ± 20
23		340 ± 38
24		600 ± 145

b



Compound	R_2	IC_{50} (nM)
25		1.1 ± 0.2
26		3.8 ± 0.7
9		5.8 ± 0.7
27		8.2 ± 0.7
28		15 ± 3
29		16 ± 1
30		27 ± 3
8		68 ± 16
31		500 ± 55
32		560 ± 130

[a]Compounds with variation at the R_1 substituent (a) and the R_2 substituent (b) are ordered according to their potency. The series also includes compounds from Table 2a for comparison.

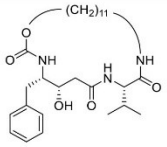
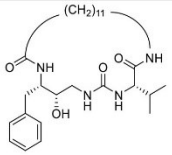
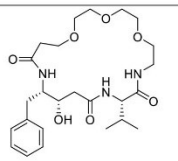
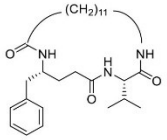
Furthermore, we selected two representative macrocycles to demonstrate that highly potent inhibitors (with single-digit nanomolar IC_{50} values) can be optimized to display desired properties including solubility and cell permeability. Ionizable **25** showed better solubility than pepstatin A (342 and 197 $\mu g \cdot mL^{-1}$, respectively), and nonionizable **33** displayed better permeability in Caco 2 cells, with moderate efficiency compared to the poor efficiency measured for pepstatin A (P_{app} of 16.8 × 10^{-7} and 1.1 × 10^{-7} $cm \cdot s^{-1}$, respectively). Notably, these representative macrocycles were noncytotoxic, similar to pepstatin A (Supporting Information Table 2). In conclusion, the developed macrocyclic scaffold displays drug-like properties and significantly improves upon the pepstatin A parent molecule in a number of parameters relevant to drug design.

Synthesis of Pepstatin A Fragments and Macrocycles. Linear truncated analogues of pepstatin A (**2**−**5**) were prepared by standard solution techniques from commercially available amino acid derivatives. During synthesis, a *t*-butoxycarbonyl (Boc) group was used for transient protection of the α-amino group, whereas the terminal carboxylate was left unprotected. *N*-hydroxysuccinimidyl esters were used for coupling reactions; these esters were purchased or generated in situ using *N*,*N*,*N*′,*N*′-tetramethyl-*O*-(*N*-succinimidyl)uronium tetrafluoroborate (TSTU).

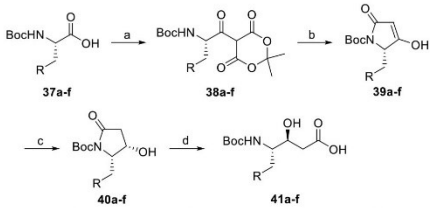
Statine analogues were prepared according to a modified version of the procedure described by Jouin et al.,[26] except for commercially available analogues. Preparation of the analogues (**41a**−**f**) is shown in Scheme 1. The long-chain ω-amino acids were purchased, with the exception of methyl 14-amino-

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

Table 4. Selected Backbone Modifications of Macrocyclic Inhibitors of Human CatD

Compound	Structure	IC_{50} (nM)
33		2.5 ± 0.6
34		21 ± 3
35		850 ± 210
36		2000 ± 340

Scheme 1. Synthesis of Statine Analogues[a]



a: R = 2-naphthyl; **b**: R = 4-thiazolyl; **c**: R = 4-Me-C_6H_4-CH_2S; **d**: R = 3,4-dichlorophenyl; **e**: R = 3,4-difluorophenyl; **f**: R = 4-fluorophenyl

[a](a): R = 2-naphthyl; (b): R = 4-thiazolyl; (c): R = 4-Me-C_6H_4-CH_2S; (d): R = 3,4-dichlorophenyl; (e): R = 3,4-difluorophenyl; (f): R = 4-fluorophenyl. Reagents and conditions: (a) DCC, DMAP, Meldrum's acid, DCM; (b) EtOAc, reflux; (c) $NaBH_4$, DCM/AcOH/10:1; (d) NaOH, acetone.

tetradecanoate (**46**) and methyl 15-aminopentadecanoate. Compound **46** was prepared from Boc-12-aminododecanoic acid (**42**) as shown in Scheme 2.[27] Methyl 15-aminopentadecanoate was prepared according to the published procedure.[28]

Macrocyclic inhibitors with an aliphatic P2−P3′ bridge (**6**−**32**) were prepared according to Scheme 3. First, a Boc-protected amino acid (**47**) was coupled with an ω-amino fatty acid methyl ester (**48**) using standard coupling conditions to yield **49**. After removal of the Boc group with TFA, the free amine was coupled with a corresponding Boc-protected statine analogue to yield **50**. Methyl ester **50** was then hydrolyzed using LiOH, and carboxylic acid **51** was converted to an *N*-hydroxysuccinimidyl ester with TSTU. The Boc-group was removed with TFA, and the intermediate was cyclized by slow addition into basic solution. Compounds **33**, **34**, and **35** were prepared by a slightly modified procedure; details are provided in the Experimental Section. A different bond was chosen for cyclization of the PEG-containing inhibitor **35** due to availability of the Boc-protected precursor (Scheme 4). A similar procedure was used to synthesize the carbamate-containing compound **33** (Scheme 5).

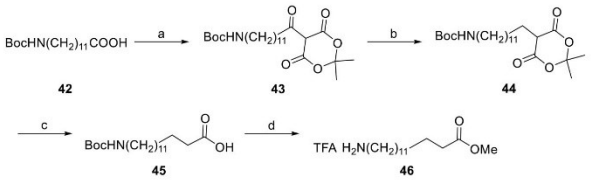
Crystal Structures of Three Macrocyclic Inhibitors in Complex with Human CatD. *Macrocyclic Inhibitors Target the Central Part of the CatD Active Site.* Native human CatD was crystallized in complexes with macrocyclic inhibitors **9**, **14**, and **33** (Figure 2a), which exhibit high inhibitory potency in the low nanomolar and subnanomolar range. The structures were determined by molecular replacement using the structure of uncomplexed CatD (PDB code 1LYA) as a model. All complexes crystallized in the orthorhombic $C222_1$ group with one molecule in the asymmetric unit. The CatD-**9**, CatD-**14**, and CatD-**33** structures were refined to resolutions of 1.55, 1.80, and 1.85 Å, respectively. The CatD structures consist of two chains: the light chain spans residues 1−97 and the heavy chain includes residues 106−347 (mature CatD numbering). One N-terminal residue from the heavy chain and one C-terminal amino acid from both chains could not be located in the electron density map. Electron densities used for modeling of inhibitors were of excellent quality in all structures (Figure 2b). The CatD structures did not show any significant differences; root-mean-square deviation for superposition of protein backbones was 0.14 Å for CatD-**14** versus CatD-**33**/CatD-**9** and 0.11 Å for CatD-**33** versus CatD-**9**. All structures are *N*-glycosylated at Asn70 and Asn199; the Asn70 glycan in CatD-**9** occurs in two conformations.

The complexed macrocycles contain valine in the P2′ position and a statine derivative with an aromatic ring (benzyl for **9** and **33**, β-naphthyl for **14**) in the P1 position. They are linked by an aminododecanoic acid bridge, which is extended by an oxygen atom in the CatD-**33** complex. Positions of the macrocyclic scaffold in all structures overlap; P1 and P2′ are buried deep inside the active site pocket, while the aliphatic bridge faces out (Figure 2a). The orientation of the β-naphthyl group in CatD-**14** differs slightly from that of aromatic rings in the other structures (Figure 2c), and there is an additional atomic contact formed between this substituent and the protein backbone (Thr125, Supporting Information Table 3). The prolongation of the bridge by one atom in CatD-**33** resulted in a change in bridge conformation associated with formation of a new hydrophobic interaction with Met307 and loss of a hydrogen bond with Ser80 compared to the other complexes (Supporting Information Table 3). In the CatD-**14** structure, a second inhibitor molecule was found outside the active site and appears to be a crystallographic artifact caused by an excess of inhibitor during crystallization.

Small Macrocycles Mimic the Binding Conformation of Pepstatin A and Establish a Similar Number of Interactions. The active site of CatD contains a catalytic dyad of two aspartic residues (Asp33 and Asp231) in a conserved Asp-Thr-

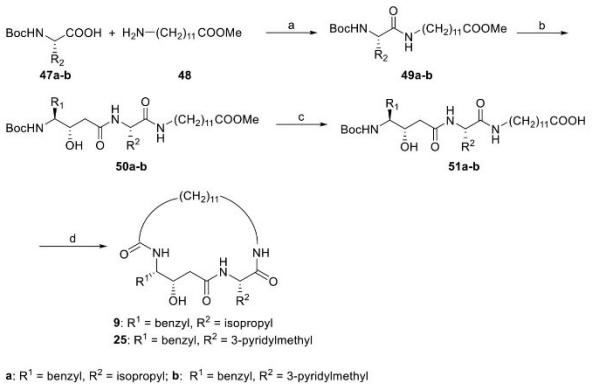
https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

Scheme 2. Synthesis of Methyl 14-Aminotetradecanoate[a]



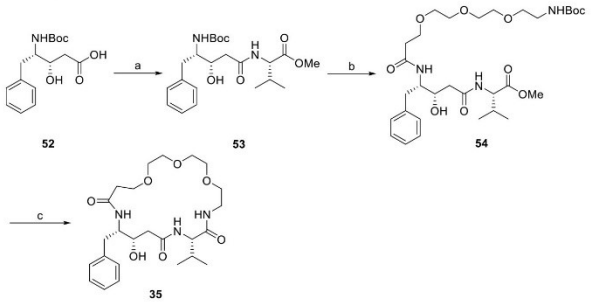
[a]Reagents and conditions: (a) DCC, DMAP, Meldrum's acid, DCM; (b) $NaBH_4$, DCM/AcOH/10:1; (c) 110 °C, DMF/H_2O/10:1; (d) (1) $SOCl_2$, MeOH, 0 °C; (2) TFA.

Scheme 3. General Synthesis of Macrocycles with an Aliphatic Chain[a]



[a](a): R^1 = benzyl, R^2 = isopropyl; (b): R^1 = benzyl, R^2 = 3-pyridylmethyl. Reagents and conditions: (a) HBTU, DIEA, DMF; (b) (1) TFA, DCM; (2) Boc-statine-OH, HBTU, DIEA, DMF; (c) LiOH, THF/H_2O; (d) (1) TSTU, DIEA, DMF; (2) TFA; (3) DIEA, DMF.

Scheme 4. Synthesis of Macrocycle with a PEG Chain[a]



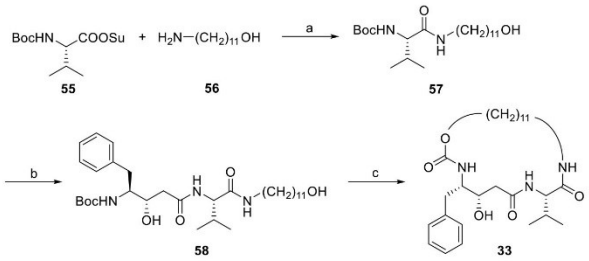
[a]Reagents and conditions: (a) (1) HBTU, DIEA, DMF; (2) H-Val-OMe; (b) (1) TFA; (2) BocHN-PEG-3-CH_2CH_2COOH, HBTU, DIEA, DMF; (c) (1) LiOH, THF/H_2O; (2) TFA; (3) DPPA, DIEA, DMF.

Gly motif. The dyad is hydrogen bonded to the hydroxyl group of the phenylstatine/naphthylstatine residue of macrocyclic inhibitors (Figure 2c), occupying the position of a water molecule that interacts with the dyad in uncomplexed CatD (PDB code 1LYA). This interaction, which mimics a tetrahedral intermediate of the enzymatic reaction, is the key inhibitory determinant in the CatD−macrocycle complexes and is identical to that observed in the CatD−pepstatin A complex (PDB code 1LYB).

The macrocyclic inhibitors occupy the S2 to S3′ subsites of CatD, which are part of the larger binding area of linear pepstatin A spanning the S4 to S4′ subsites. However, both types of inhibitors form comparable numbers of atomic contacts and hydrogen bonds (Supporting Information Table 3). The three macrocyclic inhibitors share a set of common

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

Scheme 5. Synthesis of Macrocycle with a Carbamate Moiety[a]



[a]Reagents and conditions: (a) DIEA, DMF; (b) (1) TFA, DCM; (2) Boc-Statine-OH, TSTU, DIEA, DMF; (c) (1) DSC, DMF; (2) TFA, DCM; (3) DIEA, DMF.

interactions with the CatD active site residues. These include hydrogen bonds with Asp33, Gly35, His77, Gly79, Ser80, Tyr205, Asp231, and Gly233 (Figure 2c). Pepstatin A forms a very similar network of hydrogen bonds with a few differences. There are two extra hydrogen bonds between Ser235 and the N-terminal segment of pepstatin A, whereas the hydrogen bond formed between His77 and macrocyclic inhibitors is absent in the pepstatin A complex (Figure 2e). In addition, macrocyclic inhibitors form more hydrophobic interactions than pepstatin A due to bulkier groups in the P1 and P2′ positions (Figure 2d). The benzyl and β-naphthyl aromatic rings in the P1 position of macrocycles form hydrophobic contacts with Phe126 and Phe131, whereas pepstatin A, with a smaller isobutyl group in P1, lacks these interactions. Similarly, the P2′ Val residue of macrocycles forms more contacts than the P2′ Ala of pepstatin A (Supporting Information Table 3).

Computational Analysis of the Binding Mode of Macrocycles. We evaluated the interaction energies of macrocyclic inhibitors **9**, **14**, and **33** bound to CatD subsites by performing quantum chemical calculations on the crystallographic complexes (Figure 3a). We found that the interaction of the hydroxyl group of the inhibitor phenylstatine/naphthylstatine residue has the most negative $\Delta G'$. The critical role of this interaction is reflected in the high IC_{50} measured for macrocycle derivative **36**, which lacks this hydroxyl (Table 4). Furthermore, **14** had a more negative P1 $\Delta G'$ value than **9** and **33**, which might be due to the more hydrophobic β-naphthyl group of **14** compared to phenyl. The bridge segments (br and br′) in **33** formed more favorable interactions than those in **9** and **14**. None of the segments of **9** formed considerably stronger interactions compared to the other two inhibitors, which correlates with their IC_{50} values.

Next, we computed the energy contributions of individual amino acid residues in the CatD active site interacting with the macrocyclic inhibitors in crystallographic complexes (Figure 3b). These results confirmed the critical contribution of the catalytic Asp33 and Asp231 residues. Additionally, we observed a large contribution from residues Tyr78, Ser80, and Tyr205. To gain deeper insights into the effects of P1 and P2′ variations, we modeled and scored inhibitors **8**, **22**, **24**, **25**, and **31** in the CatD active site using the crystallographic complexes as a guide (Supporting Information Figure 1). These representative inhibitors were selected so that the substituents differed markedly in size and hydrophobicity. Compounds **22** and **24**, in which the hydrophobic P1 phenyl group of **9** is replaced with less hydrophobic heterocycles, had less negative scores due to lower negative interaction energy terms, likely reflecting the missing optimal interaction partners in the S1 subsite. Compounds **25**, **8**, and **31** differ in P2′. The Val of **9** is replaced with Ala in **8**, which resulted in a less negative interaction energy term and total score. In **31** with P2′ Thr, the hydroxyl group forms only one hydrogen bond with the resident water molecule W14. The hydroxyl likely loses some hydrogen bonds upon binding, and the favorable increase in the interaction energy term is thus smaller than the unfavorable change in the desolvation term (Supporting Information Figure 1). The resulting score was less negative than that of **9**, which correlates with the IC_{50} values observed for these inhibitors. Introduction of a pyridine group in **25** decreased the IC_{50} value. Our in silico model showed that the N atom of pyridine forms a hydrogen bond with the His77 backbone, and the ring forms a stacking interaction with the Val144 side chain. The corresponding favorable change in interaction energy term is larger than the increase in the desolvation term (Supporting Information Figure 1), resulting in a more negative score.

Inhibition Kinetics of Macrocycles Differ from Those of Pepstatin A. Compound **9** was chosen as a model macrocyclic inhibitor for a detailed analysis of inhibition kinetics with human CatD. First, we demonstrated that the inhibitor has a competitive mode of inhibition, consistent with its crystallographic binding mode (Supporting Information Figure 2a). Second, we monitored the inhibition after adding enzyme over a prolonged period of time to assess the rate at which the system establishes equilibrium. A linear dependency was observed (Supporting Information Figure 2b), suggesting that the macrocyclic inhibitors do not follow slow-binding kinetics, as has been previously described for the interaction of pepstatin A with pepsin.[29] Pepstatin A exhibits tight-binding properties with aspartic proteases and is therefore used as a titrant for determination of active enzyme concentration.[18,21] However, we did not observe tight-binding properties for the macrocyclic inhibitor; its potency was independent of the enzyme concentration in the assay (Supporting Information Figure 2c). These findings suggest that redesigning pepstatin A into a small macrocyclic mimetic is associated with changes in kinetic inhibition behavior.

Macrocycles Selectively Inhibit CatD and Related Aspartic Proteases. We assessed the selectivity of macrocyclic inhibitors using a panel of model aspartic proteases from the A1 (pepsin) and A2 (retropepsin) families. At 50 nM inhibitor concentration, human and tick CatDs, yeast

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

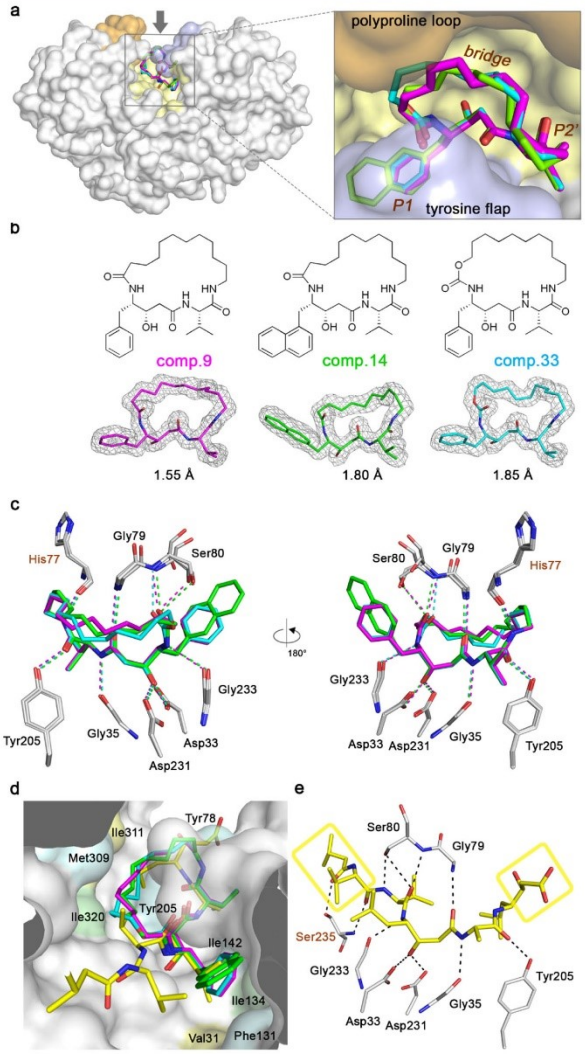


Figure 2. Crystal structures of human CatD in complex with three macrocyclic inhibitors (**9**, **14**, and **33**, PDB codes 6QCB, 6QBG, and 6QBH). (**a**) Top view of the CatD active site with the superimposed macrocyclic inhibitors. The enzyme is shown as semitransparent surface and the inhibitors as sticks (**9**, magenta; **14**, green; **33**, cyan). The active site residues are colored as follows: Tyr flap residues are pale blue, polyproline loop residues are pale orange, and the rest of the residues are light yellow. (b) Chemical structures and electron density maps of macrocyclic inhibitors. The $2F_o - F_c$ maps are contoured at 1.2σ. (c) Hydrogen bond networks (dashed lines) formed between CatD and macrocyclic inhibitors are shown in two views rotated by 180°. Inhibitors are colored as in (b); interacting enzyme residues are gray. Heteroatoms have standard color coding. (d) Superposition of macrocyclic inhibitors colored as in (b) and pepstatin A (yellow) in the CatD active site. The enzyme is in semitransparent surface representation, and residues that form hydrophobic interactions with inhibitors are highlighted. The residues interacting with the macrocycles are pale cyan, with pepstatin A are yellow, and with both inhibitors are pale green. (e) Hydrogen bond network (dashed lines) formed between CatD and pepstatin A. Pepstatin A is colored yellow; interacting enzyme residues are gray. Yellow boxes indicate segments of pepstatin A that are absent in the macrocycle scaffold. Residues labeled in brown form pepstatin A-specific hydrogen bonds; macrocycle-specific hydrogen bonding is labeled in (c).

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

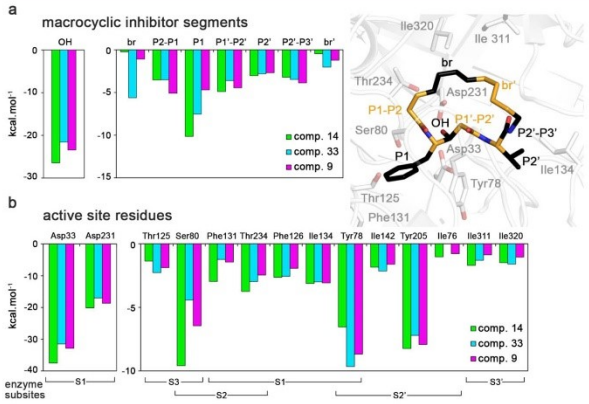


Figure 3. Quantum chemical calculations of interaction "free" energies of three crystallographically characterized macrocyclic inhibitors in the CatD active site. The 3D view shows fragmentation of an inhibitor into segments (alternating black and orange) and the active site residues of CatD (gray) used for the calculations (PDB code 6QCB). (a) Contributions of individual inhibitor segments. (b) Contributions of the active site residues obtained from a "virtual glycine" scan.

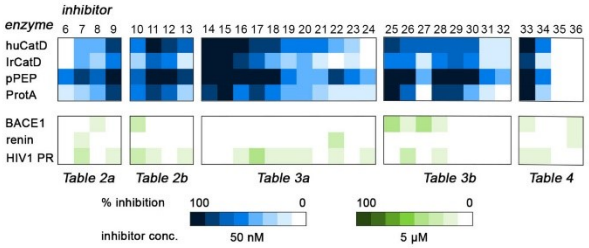


Figure 4. Inhibitory selectivity of macrocyclic compounds against representative aspartic proteases. Inhibition of proteolytic activity was screened in a kinetic activity assay with FRET fluorogenic substrates. Inhibitors were applied at 50 nM concentration (upper panel) for human cathepsin D, tick cathepsin D1 (IrCatD), yeast proteinase A (ProtA), and porcine pepsin (pPEP), and at 5 μM concentration (lower panel) for human β-secretase (BACE1), human renin, and HIV-1 protease (HIV1 PR). Data are expressed as % inhibition relative to the uninhibited control (0%) and displayed as a heat map (blue and green scales). Inhibitors are grouped according to their appearance in data tables.

proteinase A, and porcine pepsin were generally susceptible to inhibition but exhibited variations in selectivity profile (Figure 4). This is consistent with the fact that these A1-family enzymes are closely related with very similar active site structure and substrate specificity. Several of the macrocyclic compounds displayed a substantial degree of selectivity. Compounds **20**, **21**, and **27** exhibited a preference for human CatD, whereas **7**, **12**, **18**, **22**, **26**, and **28**−**30** preferred porcine pepsin. In contrast, macrocyclic inhibitors were not effective at 50 nM concentration against more distantly related aspartic proteases, including human renin and BACE1 from the A1 family and HIV-1 protease from the A2 family. Increasing the inhibitor concentration to 5 μM resulted in only modest inhibitory activity, with pronounced differences in selectivity (Figure 4).

Furthermore, we assessed the proteolytic stability of the macrocycles. Human CatD, papain, trypsin/chymotrypsin, and thermolysin were used as typical representatives of four major protease classes: aspartic proteases, cysteine proteases, serine proteases, and metalloproteases, respectively. Selected macrocycles (**25** and **33**) subjected to exhaustive treatment with these proteases did not undergo significant degradation (see Experimental Section). Notably, the non-aspartic proteases papain, trypsin/chymotrypsin, and thermolysin were not inhibited by **25** and **33**, which further demonstrates the high selectivity of the macrocycles.

In conclusion, our macrocyclic inhibitor scaffold is proteolytically stable and effectively regulates the activity of human CatD and its close homologues from the A1 family of aspartic proteases. The inhibitory selectivity can be fine-tuned by introducing structural diversity into the P1, P2′, and linker regions.

■ DISCUSSION AND CONCLUSIONS

The high affinity of naturally occurring pepstatin A for human CatD makes it an attractive template for biomimetic inhibitor design.[18] Here, we used the crystal structure of the CatD−pepstatin A complex and previously obtained SAR data as guides for rational design of small peptidomimetic inhibitors.[19,30] We aimed to avoid the inherent difficulties connected with use of linear peptides as therapeutics. Linear peptides suffer from poor oral absorption, short duration of

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

action, proteolytic instability, and rapid biliary clearance. In contrast, cyclic peptidomimetics in general have better pharmacokinetic properties and are being increasingly employed in drug development.[25,31−33]

First, we evaluated linear fragments of pepstatin A to assess size requirements for the inhibitory motif. Second, we analyzed pepstatin A bound in the active site of human CatD. The inhibitor adopts a U-shaped conformation with several side chains in positions that can be effectively connected by an aliphatic bridge to impose conformational constraint by cyclization. Based on these data, we designed a functional macrocyclic scaffold with minimized size. High-resolution crystal structures of CatD−macrocycle complexes revealed that this scaffold stabilizes the active site conformation seen in pepstatin A and forms a comparable number of atomic contacts and hydrogen bonds as the larger structure of the parent ligand. Quantum chemical calculations of the interaction energies of macrocycles demonstrated the critical role of the statine residue, which is correctly positioned in the scaffold to interact with the enzyme catalytic center. The results of the structural analysis are in line with findings that proteases recognize peptide ligands with extended β-strand conformation and that this extended-like conformation can be induced in short peptides by macrocyclization to improve their affinity.[34,35]

We prepared a library of more than 30 macrocyclic compounds with variable bridge size and various side chains in the P1 and P2′ positions. SAR analysis revealed that the minimized scaffold effectively potentiates inhibition, and we identified human CatD inhibitors with low nanomolar and subnanomolar potency. The inhibitor selectivity can be altered by residue substitutions in macrocycles to preferentially target human CatD or its close homologues from the A1 family of aspartic proteases. Notably, the developed compounds display favorable drug-like properties comparable to those reported for other medicinally relevant macrocycles.[25,31−33] They meet Lipinski's "rule of five" criteria and other physicochemical parameters for drug design that pepstatin A generally does not. Furthermore, the macrocycles are proteolytically stable and noncytotoxic, and parameters such as solubility and cell permeability can be controlled by structural substitutions. In conclusion, we present new biomimetic macrocyclics as a powerful template for further optimization and future development of potential chemotherapeutics against pathologies associated with human CatD and CatD-like proteases.

■ EXPERIMENTAL SECTION

Materials. Mature human CatD was isolated from human placenta as previously described[36] with minor modifications. Briefly, the tissue was homogenized and resuspended in 3% NaCl solution. The pH was adjusted to 3.5, and the suspension stirred for 12 h and centrifuged at 6000*g* for 1 h. The supernatant was precipitated with 45% saturated ammonium sulfate, centrifuged at 11,000*g* for 1 h, and precipitated with 65% ammonium sulfate. The pellet was resuspended in water and dialyzed against 1 mM EDTA. The resulting material was separated on a Sephadex G150 column (GE Healthcare) equilibrated with 10 mM sodium phosphate, pH 6.8. Fractions with CatD activity were pooled and dialyzed against equilibrated in 30 mM sodium acetate, pH 3.5, 0.5 M NaCl. CatD was purified by pepstatinyl-Sepharose affinity chromatography. Several wash steps (30 mM sodium acetate, pH 3.5; 30 mM sodium acetate, pH 3.5, 5% Triton X-100; and 30 mM Tris-HCl, pH 7.5, 0.5 M NaCl) were performed, and CatD was eluted with 30 mM Tris-HCl, pH 8.2, 0.5 M NaCl. The elution fraction was buffer-exchanged into 25 mM Tris-HCl, pH 7.0, and concentrated by ultrafiltration to 7.5 mg mL^{-1}.

CatD from the hard tick *Ixodes ricinus* was produced in the bacterial expression, refolded, and chromatographically purified as previously described.[37] Human BACE1 and human renin were purchased from R&D and Enzo, respectively; yeast proteinase A, porcine pepsin, bovine trypsin, bovine chymotrypsin, papain, and thermolysin were purchased from Sigma-Aldrich. HIV-1 protease and its substrate were kindly donated by Pavlína Řezáčová and Tat'ána Majerová, respectively. The substrates for BACE1 and renin were purchased from R&D and Cayman Chemicals, respectively; the substrates for non-aspartic proteases were from Bachem; the CatD substrate was prepared as previously described;[38] and pepstatin A was purchased from Sigma-Aldrich.

General Procedures. Commercially available high-performance liquid chromatography (HPLC) grade acetonitrile, catalysts, and reagent-grade materials were used for synthesis. TLC was performed on Silica gel 60 F254-coated aluminum sheets (Merck). Flash chromatography was performed on Silica gel 60 (0.040−0.063 mm, Fluka). All chemicals were purchased from Sigma-Aldrich, TCI, Combi-Blocks, or Iris Biotech GmbH and were used without further purification. ^{1}H and ^{13}C NMR spectra were measured using Bruker AVANCE III HD 400 MHz, Bruker AVANCE III HD 500 MHz, and Bruker AVANCE III 600 MHz instruments. The internal signal of TMS (δ 0.0, $CDCl_3$) or residual signal of $CDCl_3$ (δ 7.26), DMSO-d_6 (δ 2.50, 3.33), or d_4-MeOD (δ 3.31, 4.87) was used for standardization of ^{1}H NMR spectra. For ^{13}C NMR spectra (APT experiments), the residual signal of $CDCl_3$ (δ 77.16), DMSO-d_6 (δ 39.52), or d_4-MeOD (δ 49.00) was used. NMR spectra were recorded at room temperature unless noted otherwise. Chemical shifts are given in δ scale; coupling constants (J) are given in Hz. The ESI mass spectra were recorded using a ZQ micromass mass spectrometer (Waters) equipped with an ESCi multimode ion source and controlled by MassLynx software. Low-resolution ESI mass spectra were recorded using a quadrupole orthogonal acceleration time-of-flight tandem mass spectrometer (Q-Tof micro, Waters) and high-resolution ESI mass spectra using a hybrid FT mass spectrometer combining a linear ion trap MS and Orbitrap mass analyzer (LTQ Orbitrap XL, Thermo Fisher Scientific). The conditions were optimized for suitable ionization in the ESI Orbitrap source (sheath gas flow rate of 35 au, aux gas flow rate of 10 au of nitrogen, source voltage of 4.3 kV, capillary voltage of 40 V, capillary temperature of 275 °C, tube lens voltage of 155 V). The samples were dissolved in methanol and applied by direct injection. Inhibitors were purified by preparative scale HPLC on a JASCO PU-975 instrument (flow rate of 10 mL min^{-1}) equipped with a UV-975 UV detector and Waters YMC-PACK ODS-AM C18 Prep Column (5 μm, 20 × 250 mm). The purity of compounds was assessed on an analytical JASCO PU-1580 HPLC (flow rate of 1 mL min^{-1}, invariable gradient from 2 to 100% acetonitrile in 30 min) with a Watrex C18 Analytical Column (5 μm, 250 × 5 mm). The final inhibitors were of at least 95% purity. Optical rotations were measured using an Autopol IV instrument (Rudolph Research Analytical). All final compounds passed the PAINS filter using false positive remover.[39] Aqueous solubility of compounds was determined in phosphate-buffered saline, pH 7.4, at 26 °C, and quantified by HPLC-MS using a Waters ACQUITY C18 BEH column (1.7 μm, 2.1 × 100 mm) with a formic acid/acetonitrile elution system and UV detection.

*Truncated Pepstatin A Analogues (**2**−**5**). Isovaleroyl-Val-Val-Sta-Ala-OH (**2**).* ^{1}H NMR (400 MHz, CD_3OD): 0.83−1.03 (24H, m), 1.27−1.37 (2H, m), 1.40 (3H, d, J = 7.3), 1.52−1.67 (2H, m), 1.75−1.83 (1H, m), 1.99−2.16 (5H, m), 2.28−2.36 (2H, m), 2.63 (1H, p, J = 1.9), 3.09−3.17 (1H, m), 3.91−4.05 (2H, m), 4.13−4.23 (2H, m), 4.36 (1H, q, J = 7.3), 7.72 (1H, d, J = 9.0), 8.00 (1H, t, J = 8.6). ^{13}C NMR (101 MHz, CD_3OD): 17.9, 18.9, 19.1, 19.9, 20.1, 22.2, 22.8 (2C), 23.9, 25.8, 27.5, 31.6, 31.7, 41.1, 41.6, 46.0, 49.6, 52.9, 60.4, 60.6, 71.3, 173.5, 173.6, 173.8, 175.5, 176.5. ESI MS: 551 ([M + Na]$^+$). HR ESI MS: calcd for $C_{26}H_{48}O_7N_4Na$, 551.3415; found, 551.3413.

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

*Isovaleroyl-Val-Val-Sta-OH (**3**).* ^{1}H NMR (400 MHz, CD_3OD): 0.85−1.00 (24H, m), 1.26−1.39 (1H, m), 1.47−1.68 (2H, m), 1.95−2.18 (6H, m), 2.36 (1H, dd, J = 15.9, 8.7), 2.45 (1H, dd, J = 15.9, 4.3), 3.96−4.08 (2H, m), 4.16−4.23 (2H, m), 7.71 (1H, d, J = 9.1), 7.98 (1H, d, J = 8.4), 8.06 (1H, d, J = 8.7). ^{13}C NMR (101 MHz, CD_3OD): 18.9, 19.1, 19.9, 20.0, 22.3, 22.8, 23.8, 25.7, 27.5, 31.7, 31.8, 39.8, 41.1, 46.0, 52.5, 60.3, 60.4, 70.6, 173.3, 173.7, 175.3, 175.5. ESI MS: 458 ([M + H]$^+$). HR ESI MS: calcd for $C_{23}H_{44}O_6N_3$, 458.3225; found, 458.3227.

*Isovaleroyl-Sta-Ala-Sta-OH (**4**).* ^{1}H NMR (400 MHz, CD_3OD): 0.83−1.01 (18H, m), 1.26−1.41 (5H, m), 1.51−1.70 (4H, m), 2.06−2.18 (3H, m), 2.28−2.49 (4H, m), 3.97−4.06 (4H, m), 4.31 (1H, q, J = 7.2). ^{13}C NMR (101 MHz, CD_3OD): 18.2, 22.4, 22.7, 23.0, 23.7, 23.8, 25.9, 26.1, 27.3, 39.9, 41.1, 41.2, 46.4, 50.9, 52.0, 52.5, 71.0, 71.3, 174.0, 175.3, 175.4, 175.7. ESI MS: 488 ([M + H]$^+$). HR ESI MS: calcd for $C_{24}H_{46}O_7N_3$, 488.3330; found, 488.3331.

*Isovaleroyl-Sta-Ala-OH (**5**).* ^{1}H NMR (400 MHz, CD_3OD): 0.88−1.03 (12H, m), 1.25−1.38 (1H, m), 1.40 (3H, d, J = 7.3), 1.51−1.68 (2H, m), 2.02−2.17 (3H, m), 2.24−2.42 (2H, m), 3.96 (1H, ddd, J = 7.8, 5.0, 2.6), 4.01−4.11 (1H, m), 4.38 (1H, q, J = 7.3). ^{13}C NMR (101 MHz, CD_3OD): 17.6, 22.1, 22.7, 22.9, 23.9, 26.1, 27.4, 41.0, 41.5, 46.4, 49.4, 52.5, 71.6, 173.8, 175.8, 176.0. ESI MS: 329 ([M − H]$^-$). HR ESI MS: calcd for $C_{16}H_{29}O_5N_2$, 329.2082; found, 329.2078.

*Macrocyclic Inhibitors (**6**−**36**). (3S,7S,8S)-7-Hydroxy-8-isobutyl-3-methyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**6**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.80 (3H, d, J = 6.4), 0.86 (3H, d, J = 6.5), 1.10−1.64 (21H, m), 2.06−2.26 (4H, m), 2.76−2.86 (1H, m), 3.18−3.28 (1H, m), 3.72−3.82 (1H, m), 3.88 (1H, q, J = 6.7), 4.05 (1H, p, J = 7.2), 4.79 (1H, d, J = 5.3), 7.44 (1H, d, J = 9.3), 7.93 (1H, d, J = 7.0), 7.97 (1H, t, J = 5.8). ^{13}C NMR (101 MHz, DMSO-d_6): 17.6, 21.7, 23.3, 24.4, 24.9, 26.1, 26.6, 26.7, 27.1, 27.3, 27.8, 34.9, 38.3, 39.8, 48.9, 49.1, 69.5, 170.3, 172.2, 172.5. ESI MS: 448 ([M + Na]$^+$). HR ESI MS: calcd for $C_{23}H_{43}O_4N_3Na$, 448.31458; found, 448.31401.

*(3S,7S,8S)-7-Hydroxy-8-isobutyl-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**7**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.80 (3H, d, J = 6.5), 0.84 (6H, d, J = 6.8), 0.86 (3H, d, J = 6.5), 1.14−1.57 (21H, m), 1.95−2.03 (1H, m), 2.16−2.21 (2H, m), 2.13 (1H, dd, J = 14.5, 6.1), 2.21 (1H, dd, J = 14.5, 8.0), 2.77−2.85 (1H, m), 3.19−3.28 (1H, m), 3.77 (1H, dddd, J = 10.7, 10.5, 4.1, 1.6), 3.93−3.98 (1H, m), 3.86−3.90 (1H, m), 4.79 (1H, br s), 7.44 (1H, d, J = 9.4), 7.93 (1H, d, J = 6.9), 7.97 (1H, dd, J = 6.4, 5.1). ^{13}C NMR (101 MHz, DMSO-d_6): 18.0, 19.2, 21.6, 23.3, 24.4, 24.8, 24.9, 26.1, 26.6, 26.7, 27.0, 27.2, 27.8, 29.5, 34.8, 38.3, 39.7, 48.8, 49.1, 69.5, 170.3, 172.1, 172.5. ESI MS: 476 ([M + Na]$^+$). HR ESI MS: calcd for $C_{25}H_{47}O_4N_3Na$, 476.34588; found, 476.34557.

*(3S,7S,8S)-8-Benzyl-7-hydroxy-3-methyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**8**).* ^{1}H NMR (600 MHz, DMSO-d_6): 0.92−1.02 (2H, m), 1.09−1.45 (18H, m), 2.06−2.12 (3H, m), 2.16−2.26 (2H, m), 2.68−2.81 (3H, m), 3.21 (1H, dq, J = 14.0, 7.1), 3.86−3.92 (1H, m), 3.99−4.03 (1H, m), 4.07 (1H, p, J = 7.2), 5.02 (1H, br s), 7.11−7.25 (5H, m), 7.64 (1H, d, J = 9.3), 7.91−7.99 (2H, m). ^{13}C NMR (150 MHz, DMSO-d_6): 17.6, 24.8 (2C), 26.0, 26.5, 26.6, 27.0, 27.1, 27.2, 27.6, 30.7, 34.8, 36.9, 38.3, 39.9, 40.4, 48.9, 53.2, 69.4, 125.8, 127.9, 128.9, 139.5, 170.1, 172.2, 172.4. ESI MS: 482 ([M + Na]$^+$). HR ESI MS: calcd for $C_{26}H_{41}O_4N_3Na$, 482.29893; found, 482.29871.

*(3S,7S,8S)-8-Benzyl-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**9**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.84 (6H, d, J = 6.8), 0.97−1.07 (2H, m), 1.09−1.25 (12H, m), 1.26−1.45 (4H, m), 1.92−2.17 (3H, m), 2.18−2.34 (2H, m), 2.64−2.75 (2H, m), 2.80 (1H, dd, J = 13.8, 4.1), 3.29 (1H, dt, J = 13.8, 6.9), 3.51 (1H, s), 3.84−3.91 (1H, m), 3.92−4.02 (2H, m), 7.11−7.25 (5H, m), 7.60 (1H, d, J = 8.3), 7.66 (1H, d, J = 9.2), 7.96 (1H, dd, J = 7.0, 4.4). ^{13}C NMR (101 MHz, DMSO-d_6): 18.0, 19.3, 24.9, 25.1, 26.3, 26.8, 27.2, 27.3, 27.4, 27.6, 28.0, 29.0, 29.5, 34.8, 36.6, 38.2, 53.8, 58.8, 69.4, 125.8, 128.0, 129.0, 139.6, 170.5, 170.8, 172.5. ESI MS: 510 ([M + Na]$^+$). HR ESI MS: calcd for $C_{28}H_{45}O_4N_3Na$, 510.33023; found, 510.33010.

*(3S,7S,8S)-8-Benzyl-7-hydroxy-3-isopropyl-1,4,9-triazacycloicosane-2,5,10-trione (**10**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.82 (3H, d, J = 5.2), 0.84 (3H, d, J = 6.0), 0.92−1.04 (1H, m), 1.04−1.28 (11H, m), 1.29−1.45 (4H, m), 1.89−2.04 (2H, m), 2.12 (1H, dt, J = 13.2, 6.3), 2.18−2.30 (2H, m), 2.58−2.69 (1H, m), 2.76−2.89 (2H, m), 3.22−3.29 (1H, m), 3.91−4.04 (3H, m), 4.99 (1H, d, J = 5.2), 7.12−7.26 (5H, m), 7.64 (1H, d, J = 8.7), 7.68 (1H, d, J = 8.9), 7.86 (1H, dd, J = 7.2, 4.4). ^{13}C NMR (101 MHz, DMSO-d_6): 18.2, 19.3, 24.6, 26.9, 27.0, 27.3, 27.5, 27.8, 28.0, 29.4, 35.0, 35.6, 37.6, 40.0, 54.0, 58.5, 69.9, 125.8, 128.0, 128.9, 139.7, 170.6, 170.9, 172.2. ESI MS: 496 ([M + Na]$^+$). HR ESI MS: calcd for $C_{27}H_{43}O_4N_3Na$, 496.31458; found, 496.31396.

*(3S,7S,8S)-8-Benzyl-7-hydroxy-3-isopropyl-1,4,9-triazacyclodocosane-2,5,10-trione (**11**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.77−0.88 (6H, m), 1.03−1.52 (20H, m), 1.92−2.03 (2H, m), 2.07−2.16 (1H, m), 2.18−2.30 (2H, m), 2.60−2.70 (1H, m), 2.74−2.88 (2H, m), 3.23−3.33 (1H, m), 3.91−3.99 (2H, m), 4.04 (1H, dd, J = 8.7, 6.3), 7.11−7.29 (5H, m), 7.58 (1H, d, J = 8.7), 7.64 (1H, d, J = 9.0), 7.91 (1H, dd, J = 7.1, 4.5). ^{13}C NMR (101 MHz, DMSO-d_6): 18.0, 19.3, 24.9, 25.0, 27.0 (2C), 27.2, 27.3, 27.4, 27.6, 28.3, 29.8, 35.2, 36.1, 38.1, 40.0, 53.8, 58.1, 68.9, 125.8, 128.0, 128.9, 139.6, 170.5, 170.9, 172.3. ESI MS: 524 ([M + Na]$^+$). HR ESI MS: calcd for $C_{29}H_{47}O_4N_3Na$, 524.34588; found, 524.34564.

*(3S,7S,8S)-8-Benzyl-7-hydroxy-3-isopropyl-1,4,9-triazacyclotricosane-2,5,10-trione (**12**).* ^{1}H NMR (400 MHz, CD_3OD): 0.94 (3H, d, J = 6.9), 0.95 (3H, d, J = 6.9), 1.11−1.34 (18H, m), 1.40−1.52 (4H, m), 2.08−2.23 (3H, m), 2.37−2.49 (2H, m), 2.78−2.94 (3H, m), 3.41 (1H, dt, J = 14.0, 7.0), 4.03 (1H, ddd, J = 10.0, 4.9, 1.7), 4.11−4.17 (2H, m), 7.13−7.25 (5H, m). ^{13}C NMR (101 MHz, CD_3OD): 16.8, 18.3, 25.2, 25.7, 26.8, 27.0, 27.2, 27.4, 27.5, 27.8, 27.9, 28.1, 28.6, 29.9, 35.6, 37.5, 39.0, 40.2, 54.1, 59.4, 69.9, 125.9, 127.9, 128.9, 138.9, 172.4 (2C), 174.6. ESI MS: 538 ([M + Na]$^+$). HR ESI MS: calcd for $C_{30}H_{49}O_4N_3Na$, 538.36153; found, 538.36136.

*(3S,7S,8S)-8-Benzyl-7-hydroxy-3-isopropyl-1,4,9-triazacyclotetracosane-2,5,10-trione (**13**).* ^{1}H NMR (400 MHz, CD_3OD): 0.93 (3H, d, J = 7.1), 0.95 (3H, d, J = 7.1), 1.11−1.35 (20H, m), 1.39−1.49 (4H, m), 2.05−2.24 (3H, m), 2.38−2.52 (2H, m), 2.78−3.00 (3H, m), 3.34−3.36 (1H, m), 4.04−4.09 (1H, m), 4.12−4.21 (2H, m), 7.12−7.25 (5H, m). ^{13}C NMR (101 MHz, CD_3OD): 18.0, 19.7, 26.9, 27.4, 27.9, 28.3, 28.5, 28.6, 28.7, 29.0, 29.2, 29.6, 29.7, 30.3, 31.3, 37.1, 38.9, 40.4, 41.5, 55.4, 60.5, 71.2, 127.3, 129.3, 130.3, 140.3, 173.7, 173.8, 176.0. ESI MS: 552 ([M + Na]$^+$). HR ESI MS: calcd for $C_{31}H_{51}O_4N_3Na$, 552.37718; found, 552.37729.

*(3S,7S,8S)-7-Hydroxy-3-isopropyl-8-(naphthalen-2-ylmethyl)-1,4,9-triazacyclohenicosane-2,5,10-trione (**14**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.82 (3H, d, J = 6.7), 0.83 (3H, d, J = 6.7), 0.94−1.41 (18H, m), 1.93−2.03 (1H, m), 2.03−2.16 (2H, m), 2.25 (1H, dd, J = 14.0, 6.1), 2.33 (1H, dd, J = 14.0, 7.8), 2.61−2.68 (1H, m), 2.88 (1H, dd, J = 13.6, 10.3), 2.95 (1H, dd, J = 12.6, 4.2), 3.24−3.32 (1H, m), 3.93 (1H, dd, J = 8.0, 6.2), 3.93−3.98 (1H, m), 4.03−4.08 (1H, m), 5.10 (1H, br s), 7.38 (1H, dd, J = 8.4, 1.5), 7.39−7.47 (2H, m), 7.59 (1H, d, J = 8.0), 7.70 (1H, br s), 7.71−7.87 (4H, m), 7.98 (1H, dd, J = 7.0, 4.4). ^{13}C NMR (101 MHz, DMSO-d_6): 17.9, 19.2, 24.6, 24.8, 25.7, 26.4, 26.7, 27.0, 27.2, 27.7, 29.4, 34.8, 37.2, 38.2, 40.2, 53.7, 58.8, 69.8, 125.0, 125.7, 127.1, 127.2, 127.3, 127.4, 127.8, 131.6, 133.0, 137.2, 170.4, 170.7, 172.5. ESI MS: 560 ([M + Na]$^+$). HR ESI MS: calcd for $C_{32}H_{47}O_4N_3Na$, 560.34588; found, 560.34551.

*(3S,7S,8S)-8-(3,4-Difluorobenzyl)-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**15**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.83 (6H, d, J = 6.8), 0.88−1.48 (18H, m), 1.95−2.05 (1H, m), 2.09−2.19 (2H, m), 2.21 (1H, dd, J = 14.0, 5.8), 2.32 (1H, dd, J = 14.0, 8.3), 2.62−2.70 (1H, m), 2.70−2.79 (2H, m), 3.24−3.32 (1H, m), 3.75−3.83 (1H, m), 3.93 (1H, dd, J = 7.9, 5.9), 3.98−4.03 (1H, m), 5.11 (1H, d, J = 5.2), 7.01−7.06 (1H, m), 7.21−7.30 (2H, m), 7.60 (1H, d, J = 7.9), 7.70 (1H, d, J = 9.4), 7.98 (1H, dd, J = 7.0, 4.4). ^{13}C NMR (101 MHz, DMSO-d_6): 17.9, 19.2, 24.6, 24.8, 25.6, 26.3, 26.5, 27.0, 27.2, 27.7, 29.3, 34.7, 36.3, 38.2, 40.3, 53.5, 58.8, 69.9, 116.8 (d, J_{CCF} = 17), 117.3 (dd, J_{CCF} = 135.6, J_{CCCF} = 16.4), 125.7 (dd, J_{CCCF} = 6, J_{CCCCF} = 3.2), 137.4 (dd, J_{CCCF} = 5.8, J_{CCCCF} = 3.2), 147.9 (dd, J_{CF} = 243, J_{CCF} = 12.6), 148.9 (dd, J_{CF} = 244.3, J_{CCF} = 12.6), 170.3, 170.7, 172.6. ESI MS: 546 ([M + Na]$^+$). HR ESI MS: calcd for $C_{28}H_{43}O_4N_3F_2Na$, 546.31138; found, 546.31104.

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

*(3S,7S,8S)-8-(3,4-Dichlorobenzyl)-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**16**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.83 (6H, d, J = 6.8), 0.84−1.49 (18H, m), 1.95−2.06 (1H, m), 2.08−2.20 (2H, m), 2.20 (1H, dd, J = 14.0, 5.5), 2.33 (1H, dd, J = 14.0, 8.6), 2.61−2.68 (1H, m), 2.69−2.79 (2H, m), 3.25−3.34 (1H, m), 3.79 (1H, dddd, J = 9.7, 9.7, 4.6, 1.2), 3.92 (1H, dd, J = 7.8, 5.8), 3.99−4.04 (1H, m), 5.12 (1H, br s), 7.20 (1H, dd, J = 8.3, 1.9), 7.46 (1H, d, J = 8.3), 7.47 (1H, d, J = 1.9), 7.61 (1H, d, J = 7.9), 7.72 (1H, d, J = 9.4), 7.98 (1H, dd, J = 7.1, 4.5). ^{13}C NMR (101 MHz, DMSO-d_6): 17.8, 19.2, 24.5, 24.7, 25.6, 26.2, 26.4, 26.9, 27.0, 27.1, 27.6, 29.3, 34.7, 36.4, 38.2, 40.3, 53.4, 58.9, 70.0, 128.6, 129.5, 130.0, 130.5, 131.2, 140.9, 170.3, 170.6, 172.6. ESI MS: 578 ([M + Na]$^+$). HR ESI MS: calcd for $C_{28}H_{43}O_4N_3Cl_2Na$, 578.25228; found, 578.25196.

*(3S,7S,8S)-8-(4-Fluorobenzyl)-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**17**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.84 (6H, d, J = 6.8), 0.94−1.47 (18H, m), 1.95−2.03 (1H, m), 2.03−2.10 (1H, m), 2.10−2.17 (1H, m), 2.22 (1H, dd, J = 14.0, 6.4), 2.29 (1H, dd, J = 14.0, 7.5), 2.65−2.72 (2H, m), 2.77 (1H, dd, J = 13.7, 4.0), 3.25−3.34 (1H, m), 3.82 (1H, dddd, J = 11.0, 10.9, 3.8, 1.7), 3.94 (1H, dd, J = 8.0, 6.2), 3.96−4.01 (1H, m), 5.05 (1H, br s), 7.00−7.06 (2H, m), 7.20−7.25 (2H, m), 7.60 (1H, d, J = 8.1), 7.67 (1H, d, J = 9.2), 7.97 (1H, dd, J = 7.0, 4.4). ^{13}C NMR (101 MHz, DMSO-d_6): 17.9, 19.2, 24.8, 25.0, 26.0, 26.6, 26.9, 27.1, 27.2, 27.4, 27.8, 29.4, 34.7, 36.0, 38.1, 40.1, 53.8, 58.7, 69.5, 114.6 (d, J_{CCF} = 20.9), 130.7 (d, J_{CCCF} = 8), 135.7 (d, J_{CCCCF} = 3), 160.7 (d, J_{CF} = 240.9), 170.4, 170.7, 172.5. ESI MS: 528 ([M + Na]$^+$). HR ESI MS: calcd for $C_{28}H_{44}O_4N_3FNa$, 528.32081; found, 528.32039.

*(3S,7S,8S)-8-(Cyclohexylmethyl)-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**18**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.84 (3H, d, J = 6.8), 0.85 (3H, d, J = 6.8), 1.06−1.73 (31H, m), 1.97−2.05 (1H, m), 2.15 (1H, dd, J = 14.2, 6.3), 2.19 (2H, dd, J = 6.3, 6.1), 2.28 (1H, dd, J = 14.2, 7.8), 2.71−2.79 (1H, m), 3.28−3.37 (1H, m), 3.74−3.81 (1H, m), 3.93 (1H, dd, J = 8.0, 5.9), 3.82−3.87 (1H, m), 4.80 (1H, d, J = 5.5), 7.45 (1H, d, J = 9.2), 8.02 (1H, dd, J = 7.1, 4.4). ^{13}C NMR (101 MHz, DMSO-d_6): 17.9, 19.2, 24.9, 25.0, 25.8, 26.0, 26.1, 26.6, 27.0, 27.1, 27.2, 27.4, 27.9, 29.4, 31.9, 33.6, 33.9, 34.8, 38.2, 48.6, 58.7, 69.8, 170.6, 170.7, 172.5. ESI MS: 516 ([M + Na]$^+$). HR ESI MS: calcd for $C_{28}H_{52}O_4N_3$, 494.39523; found, 494.39501.

*(3S,7S,8R)-8-((Benzylthio)methyl)-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**19**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.85 (3H, d, J = 6.7), 0.86 (3H, d, J = 6.7), 1.15−1.62 (18H, m), 1.94−2.05 (1H, m), 2.13−2.24 (4H, m), 2.48−2.55 (1H, m), 2.59 (1H, dd, J = 13.4, 5.9), 2.78−2.85 (1H, m), 3.27−3.37 (1H, m), 3.69 (2H, s), 3.93−3.97 (1H, m), 3.98 (1H, dd, J = 8.1, 6.6), 4.06−4.11 (1H, m), 4.99 (1H, br s), 7.21−7.26 (1H, m), 7.28−7.31 (4H, m), 7.63 (1H, d, J = 9.2), 7.66 (1H, d, J = 8.2), 7.92 (1H, dd, J = 7.0, 4.5). ^{13}C NMR (101 MHz, DMSO-d_6): 18.0, 19.2, 24.9, 25.2, 26.6, 27.2, 27.4, 27.5, 27.9, 28.1, 29.5, 31.9, 34.7, 34.8, 38.1, 39.8, 51.0, 58.6, 68.0, 126.7, 128.3, 128.8, 138.5, 170.4, 170.8, 172.7. ESI MS: 556 ([M + Na]$^+$). HR ESI MS: calcd for $C_{29}H_{47}O_4N_3NaS$, 556.31795; found, 556.31774.

*(3S,7S,8S)-7-Hydroxy-8-(4-hydroxybenzyl)-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**20**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.84 (6H, d, J = 6.9), 1.02−1.46 (18H, m), 1.94−2.03 (1H, m), 2.03−2.10 (1H, m), 2.10−2.17 (1H, m), 2.21 (1H, dd, J = 14.1, 6.8), 2.26 (1H, dd, J = 14.1, 6.9), 2.57 (1H, dd, J = 13.9, 10.2), 2.69−2.77 (2H, m), 3.26−3.34 (1H, m), 3.81 (1H, dddd, J = 9.7, 9.5, 4.7, 1.6), 3.92−3.98 (3H, m), 6.59−6.63 (2H, m), 6.95−7.00 (2H, m), 7.58 (1H, br s), 7.60 (1H, d, J = 1.8), 7.96 (1H, dd, J = 7.0, 4.4), 9.05 (1H, br s). ^{13}C NMR (101 MHz, DMSO-d_6): 18.0, 19.2, 24.8, 25.1, 26.2, 26.9, 27.2, 27.3, 27.4, 27.6, 28.0, 29.5, 34.9, 35.7, 38.1, 40.1, 53.9, 58.7, 69.2, 114.7, 129.5, 129.7, 155.4, 170.5, 170.8, 172.4. ESI MS: 526 ([M + Na]$^+$). HR ESI MS: calcd for $C_{28}H_{45}O_5N_3Na$, 526.32514; found, 526.32546.

*(3S,7S,8S)-8-(4-(Benzyloxy)benzyl)-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**21**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.84 (6H, d, J = 6.8), 0.98−1.48 (18H, m), 1.94−2.02 (1H, m), 2.04−2.17 (2H, m), 2.21 (1H, dd, J = 14.1, 6.6), 2.27 (1H, dd, J = 14.1, 7.2), 2.63 (1H, dd, J = 13.7, 10.3), 2.64−2.75 (2H, m), 3.25−3.32 (1H, m), 3.78−3.84 (1H, m), 3.97 (1H, dd, J = 8.1, 6.2), 3.94−4.00 (1H, m), 5.01 (1H, s), 6.85−6.88 (2H, m), 7.09−7.14 (2H, m), 7.30−7.34 (1H, m), 7.36−7.44 (4H, m), 7.58 (1H, d, J = 8.2), 7.63 (1H, d, J = 9.2), 7.97 (1H, dd, J = 7.0, 4.4). ^{13}C NMR (101 MHz, DMSO-d_6): 18.0, 19.2, 24.8, 25.0, 26.1, 26.4, 26.7, 27.0, 27.2, 27.3, 27.5, 27.9, 29.4, 34.8, 35.9, 38.1, 53.9, 58.7, 69.1, 69.4, 114.2, 127.5, 127.7, 128.4, 129.9, 131.7, 137.2, 156.6, 170.4, 170.7, 172.5. ESI MS: 616 ([M + Na]$^+$). HR ESI MS: calcd for $C_{35}H_{51}O_5N_3Na$, 616.37209; found, 616.37168.

*(3S,7S,8S)-8-((1H-Indol-3-yl)methyl)-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**22**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.83 (6H, d, J = 6.8), 1.08−1.55 (18H, m), 1.92−2.01 (1H, m), 2.06 (1H, ddd, J = 13.8, 7.2, 4.7), 2.16 (1H, ddd, J = 13.8, 8.7, 4.8), 2.23 (1H, dd, J = 14.2, 6.2), 2.28 (1H, dd, J = 14.2, 7.4), 2.65−2.72 (1H, m), 2.75 (1H, dd, J = 14.8, 8.9), 2.96 (1H, dd, J = 14.8, 5.4), 3.22−3.30 (1H, m), 3.95 (1H, dd, J = 8.4, 6.7), 4.00−4.05 (1H, m), 5.01 (1H, br s), 6.94 (1H, ddd, J = 9.6, 8.7, 0.9), 7.03 (1H, ddd, J = 9.0, 8.0, 1.0), 7.08 (1H, d, J = 2.2), 7.25−7.35 (1H, m), 7.54 (1H, d, J = 8.0), 7.59 (1H, d, J = 8.6), 7.62 (1H, d, J = 9.0), 7.88 (1H, dd, J = 7.0, 4.6). ^{13}C NMR (101 MHz, DMSO-d_6): 18.1, 19.2, 24.9, 25.2, 26.3, 26.6, 27.1, 27.4, 27.6, 27.9, 28.0, 29.5, 34.9, 38.0, 39.7, 52.9, 58.6, 68.7, 111.2, 111.7, 118.0, 118.3, 119.8, 120.7, 122.0, 122.9, 127.5, 128.9, 136.2, 170.6, 170.8, 172.4. ESI MS: 549 ([M + Na]$^+$). HR ESI MS: calcd for $C_{30}H_{46}O_4N_4Na$, 549.34113; found, 549.34082.

*(3S,7S,8R)-7-Hydroxy-3-isopropyl-8-(mercaptomethyl)-1,4,9-triazacyclohenicosane-2,5,10-trione (**23**).* ^{1}H NMR (400 MHz DMSO-d_6): 0.84 (3H, d, J = 7.0), 0.85 (3H, d, J = 7.0), 1.15−1.62 (18H, m), 1.92−2.03 (1H, m), 2.10−2.25 (5H, m), 2.48−2.55 (1H, m), 2.67 (1H, ddd, J = 13.7, 9.1, 5.5), 2.78−2.86 (1H, m), 3.27−3.34 (1H, m), 3.72 (1H, dddd, J = 8.7, 8.6, 5.7, 2.1), 3.98 (1H, dd, J = 7.6, 7.4), 4.07−4.13 (1H, m), 4.99 (1H, br s), 7.60 (1H, d, J = 9.1), 7.62 (1H, d, J = 8.7), 7.88 (1H, dd, J = 6.6, 4.7). ^{13}C NMR (101 MHz, DMSO-d_6): 18.2, 19.2, 24.7, 25.0, 25.3, 26.8, 27.3, 27.5, 27.6, 27.7, 28.0, 28.2, 29.6, 34.9, 38.1, 39.6, 55.5, 58.5, 67.9, 170.3, 170.8, 172.8. ESI MS: 442 ([M − H]$^-$). HR ESI MS: calcd for $C_{22}H_{40}O_4N_3S$, 442.27450; found, 442.27416.

*(3S,7S,8S)-7-Hydroxy-3-isopropyl-8-(thiazol-4-ylmethyl)-1,4,9-triazacyclohenicosane-2,5,10-trione (**24**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.84 (3H, d, J = 6.8), 0.85 (3H, d, J = 6.8), 1.03−1.48 (18H, m), 1.96−2.04 (1H, m), 2.04−2.10 (1H, m), 2.10−2.17 (1H, m), 2.23 (1H, dd, J = 14.2, 6.8), 2.29 (1H, dd, J = 14.2, 7.0), 2.73−2.81 (1H, m), 2.92 (1H, dd, J = 14.8, 10.0), 2.99 (1H, dd, J = 14.8, 4.4), 3.27−3.38 (1H, m), 4.09 (1H, dddd, J = 10.0, 9.8, 4.4, 2.0), 3.96 (1H, dd, J = 8.2, 6.2), 3.99 (1H, ddd, J = 6.9, 6.8, 2.0), 5.05 (1H, br s), 7.25 (1H, d, J = 2.0), 7.62 (1H, d, J = 8.2), 7.65 (1H, d, J = 9.5), 7.96 (1H, dd, J = 7.3, 4.4), 8.97 (1H, d, J = 2.0). ^{13}C NMR (101 MHz, DMSO-d_6): 18.0, 19.2, 24.7, 25.1, 26.2, 26.8, 27.2, 27.3, 27.5, 28.0, 29.5, 32.5, 34.8, 38.1, 39.8, 51.7, 58.7, 69.3, 114.3, 153.1, 154.9, 170.4, 170.7, 172.5. ESI MS: 517 ([M + Na]$^+$). HR ESI MS: calcd for $C_{25}H_{42}O_4N_4NaS$, 517.28190; found, 517.28161.

*(3S,7S,8S)-8-Benzyl-7-hydroxy-3-(pyridin-3-ylmethyl)-1,4,9-triazacyclohenicosane-2,5,10-trione (**25**).* ^{1}H NMR (400 MHz, CD_3OD): 0.98−1.07 (2H, m), 1.12−1.32 (14H, m), 1.35−1.53 (4H, m), 2.10−2.22 (2H, m), 2.26−2.33 (2H, m), 2.75−2.89 (3H, m), 3.06−3.14 (1H, m), 3.34−3.41 (2H, m), 3.94−3.99 (1H, m), 4.05 (1H, td, J = 7.1, 1.8), 4.64 (1H, dd, J = 9.3, 5.4), 7.11−7.16 (1H, m), 7.17−7.24 (4H, m), 7.65 (1H, d, J = 9.3), 7.92 (1H, dd, J = 8.1, 5.6), 8.33−8.37 (1H, m), 8.42−8.47 (1H, m), 8.69 (1H, d, J = 5.6), 8.72 (1H, s). ^{13}C NMR (101 MHz, CD_3OD): 26.6, 27.9, 28.5, 28.6, 28.8, 29.0, 29.1, 29.2, 35.4, 36.7, 39.0, 40.7, 40.8, 41.7, 55.6, 55.8, 71.5, 127.6, 128.1, 129.6, 130.4, 140.0, 140.5, 142.0, 144.2, 147.9, 172.5, 173.8, 176.2. ESI MS: 559 ([M + Na]$^+$). HR ESI MS: calcd for $C_{31}H_{45}O_4N_4Na$, 559.32548; found, 559.32554.

*(3S,7S,8S)-8-Benzyl-3-((R)-s-butyl)-7-hydroxy-1,4,9-triazacyclohenicosane-2,5,10-trione (**26**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.76−0.84 (6H, m), 0.99−1.47 (20H, m), 1.66−1.76 (1H, m), 1.99−2.17 (2H, m), 2.18−2.28 (2H, m), 2.62−2.73 (2H, m), 2.81 (1H, dd, J = 13.8, 4.3), 3.25−3.31 (1H, m), 3.83−3.93 (1H, m), 3.93−4.02

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

(2H, m), 5.02 (1H, d, J = 5.3), 7.10−7.26 (5H, m), 7.60 (1H, d, J = 8.2), 7.65 (1H, d, J = 9.1), 7.95 (1H, dd, J = 7.3, 4.3). ^{13}C NMR (101 MHz, DMSO-d_6): 11.4, 15.5, 24.4, 24.8, 25.1, 26.3, 26.8, 27.2, 27.3, 27.4, 27.7, 28.0, 34.8, 35.9, 36.5, 38.1, 40.0, 53.8, 57.7, 69.4, 125.8, 128.0, 128.9, 139.6, 170.4, 170.8, 172.4. ESI MS: 524 ([M + Na]$^+$). HR ESI MS: calcd for $C_{29}H_{47}O_4N_3Na$, 524.34588; found, 524.34528.

*(3S,7S,8S)-8-Benzyl-3-cyclohexyl-7-hydroxy-1,4,9-triazacyclohenicosane-2,5,10-trione (**27**).* NMR spectra for this compound were recorded at 80 °C, due to limited solubility of the compound. ^{1}H NMR (600 MHz, acetic acid-d_4): 1.00−1.87 (31H, m), 2.13−2.19 (1H, m), 2.22−2.33 (1H, m), 2.50−2.63 (1H, m), 2.85−3.03 (3H, m), 3.44−3.63 (1H, m), 4.12−4.40 (2H, m), 7.16−7.36 (5H, m). ^{13}C NMR (150 MHz, acetic acid-d_4): 11.4, 15.5, 24.4, 26.5, 26.8, 26.9 (2C), 27.1, 28.6, 28.8, 29.1, 29.2, 29.4, 29.5, 29.7, 30.6, 36.8, 38.6, 40.4, 42.1, 56.1, 60.4, 71.4, 127.6, 129.6, 130.3, 139.6, 171.5, 173.8, 173.9. ESI MS: 550 ([M + Na]$^+$). HR ESI MS: calcd for $C_{31}H_{49}O_4N_3Na$, 550.36153; found, 550.36104.

*(3S,7S,8S)-3,8-Dibenzyl-7-hydroxy-1,4,9-triazacyclohenicosane-2,5,10-trione (**28**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.91−1.05 (2H, m), 1.01−1.50 (16H, m), 2.04−2.14 (3H, m), 2.19 (1H, dd, J = 14.4, 7.7), 2.64−2.80 (4H, m), 3.05 (1H, dd, J = 14.0, 4.6), 3.18−3.28 (1H, m), 3.83−3.97 (2H, m), 4.28−4.36 (1H, m), 4.97 (1H, d, J = 5.1), 7.08−7.30 (10H, m), 7.65 (1H, d, J = 9.2), 7.96 (1H, d, J = 8.2), 7.98−8.06 (1H, m). ^{13}C NMR (101 MHz, DMSO-d_6): 24.8, 24.9, 26.1, 26.6, 26.7, 27.0, 27.1, 27.3, 27.7, 34.9, 36.9, 37.0, 38.4, 40.0, 53.3, 54.6, 69.2, 125.8, 126.2, 128.1, 128.9, 129.0, 138.4, 139.5, 170.3, 170.8, 172.5. ESI MS: 558 ([M + Na]$^+$). HR ESI MS: calcd for $C_{32}H_{45}O_4N_3Na$, 558.33023; found, 558.33006.

*N-(4-((3S,7S,8S)-8-Benzyl-7-hydroxy-2,5,10-trioxo-1,4,9-triazacyclohenicosan-3-yl)butyl)acetamide (**29**).* ^{1}H NMR (500 MHz, DMSO-d_6): 0.92−1.02 (2H, m), 1.04−1.50 (20H, m), 1.58−1.69 (1H, m), 1.77 (3H, s), 2.05−2.14 (2H, m), 2.17−2.26 (2H, m), 2.36 (1H, p, J = 1.8), 2.63 (1H, p, J = 1.8), 2.66−2.80 (3H, m), 2.97 (2H, q, J = 6.6), 3.83−3.90 (1H, m), 3.95−4.02 (2H, m), 5.07 (1H, d, J = 5.1), 7.11−7.24 (5H, m), 7.68 (1H, d, J = 9.3), 7.80 (1H, t, J = 5.6), 7.85 (1H, d, J = 7.5), 7.99 (1H, t, J = 5.9). ^{13}C NMR (125 MHz, DMSO-d_6): 22.7, 23.1, 24.9, 26.0, 26.6, 26.7, 27.0, 27.1, 27.4, 27.7, 28.9, 31.1, 34.8, 36.9, 38.3, 38.4, 40.0, 53.4, 53.5, 69.5, 125.8, 128.0, 129.0, 139.6, 170.0, 170.5, 171.6, 172.4. ESI MS: 581 ([M + Na]$^+$). HR ESI MS: calcd for $C_{31}H_{50}O_5N_4Na$, 581.36734; found, 581.36685.

*(3S,7S,8S)-8-Benzyl-7-hydroxy-3-phenyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**30**).* ^{1}H NMR (500 MHz, DMSO-d_6): 1.03−1.46 (16H, m), 1.99−2.08 (1H, m), 2.10−2.18 (1H, m), 2.29 (2H, dd, J = 6.9, 2.8), 2.36 (1H, p, J = 1.8), 2.63 (1H, p, J = 1.9), 2.68 (1H, dd, J = 13.7, 10.1), 2.71−2.78 (2H, m), 2.81 (1H, dd, J = 13.8, 4.4), 3.88−3.94 (1H, m), 3.95−4.00 (1H, m), 5.18 (1H, br s), 5.22 (1H, d, J = 7.2), 7.12−7.41 (10H, m), 7.72 (1H, d, J = 9.1), 8.37 (1H, d, J = 7.2), 8.41 (1H, dd, J = 7.9, 4.6). ^{13}C NMR (125 MHz, DMSO-d_6): 24.3, 24.8, 25.1, 26.9, 27.3, 27.4, 27.6, 27.9, 28.0, 34.7, 36.5, 38.3, 40.0, 53.8, 57.8, 69.3, 125.8, 127.6, 128.0, 128.3, 129.0, 138.3, 139.7, 169.9, 170.3, 172.4. ESI MS: 522 ([M + H]$^+$). HR ESI MS: calcd for $C_{31}H_{44}O_4N_3$, 522.33263; found, 522.33248.

*(3S,7S,8S)-8-Benzyl-7-hydroxy-3-((R)-1-hydroxyethyl)-1,4,9-triazacyclohenicosane-2,5,10-trione (**31**).* ^{1}H NMR (500 MHz, DMSO-d_6): 1.00 (3H, d, J = 6.4), 1.04−1.47 (16H, m), 2.10 (2H, t, J = 6.5), 2.23 (1H, dd, J = 14.1, 6.0), 2.30 (1H, dd, J = 14.2, 8.1), 2.34 (1H, p, J = 1.9), 2.62 (1H, p, J = 1.9), 2.67−2.85 (3H, m), 3.07−3.17 (1H, m), 3.82−3.88 (1H, m), 3.94−4.01 (2H, m), 4.02−4.04 (1H, m), 4.76 (1H, d, J = 5.6), 5.06 (1H, d, J = 5.2), 7.08−7.23 (5H, m), 7.48 (1H, d, J = 8.5), 7.69 (1H, d, J = 9.2), 7.88 (1H, t, J = 5.8). ^{13}C NMR (125 MHz, DMSO-d_6): 20.6, 24.8, 25.0, 25.9, 26.4, 26.5, 26.9, 27.2 (2C), 27.4, 34.9, 37.2, 38.6, 39.9, 53.4, 59.0, 65.8, 69.7, 125.9, 128.0, 129.0, 139.5, 170.3, 170.8, 172.6. ESI MS: 512 ([M + Na]$^+$). HR ESI MS: calcd for $C_{27}H_{43}O_5N_3Na$, 512.30949; found, 512.30878.

*(3S,7S,8S)-8-Benzyl-3-((R)-1-(benzyloxy)ethyl)-7-hydroxy-1,4,9-triazacyclohenicosane-2,5,10-trione (**32**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.88−1.00 (3H, m), 1.06−1.48 (19H, m), 1.40−1.54 (1H, m), 2.13 (2H, t, J = 6.4), 2.24 (1H, dd, J = 14.1, 8.7), 2.38 (1H, dd, J = 14.1, 8.7), 2.64−2.80 (2H, m), 3.19−3.32 (1H, m), 3.82−3.90 (1H, m), 3.99−4.05 (1H, m), 4.09 (1H, qd, J = 6.3, 2.6), 4.16 (1H, dd, J = 8.7, 2.6), 4.39 (1H, d, J = 11.8), 4.50 (1H, d, J = 11.9), 7.09−7.35 (10H, m), 7.56 (1H, d, J = 8.7), 7.73 (1H, d, J = 9.3), 8.13 (1H, t, J = 5.8). ^{13}C NMR (101 MHz, DMSO-d_6): 16.8, 24.6, 24.7, 25.5, 26.0, 26.4, 26.8, 27.0, 27.1, 27.3, 34.9, 37.5, 40.0, 40.6, 53.3, 58.0, 70.1, 70.7, 74.4, 125.8, 127.3, 127.9 (2C), 129.0, 138.6, 139.5, 170.0, 170.9, 172.8. ESI MS: 602 ([M + Na]$^+$). HR ESI MS: calcd for $C_{34}H_{49}O_4N_3Na$, 602.35644; found, 602.35600.

*(4S,5S,9S)-4-Benzyl-5-hydroxy-9-isopropyl-1-oxa-3,8,11-triazacyclodocosane-2,7,10-trione (**33**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.83 (3H, d, J = 6.5), 0.85 (3H, d, J = 6.5), 1.09−1.53 (16H, m), 1.85−1.99 (1H, m), 2.19−2.30 (2H, m), 2.57 (1H, dd, J = 13.8, 10.6), 2.72−2.90 (3H, m), 3.23−3.26 (2H, m), 3.59−3.76 (2H, m), 3.94−4.01 (2H, m), 4.10 (1H, dt, J = 11.4, 5.9), 6.92 (1H, d, J = 9.1), 7.11−7.29 (5H, m), 7.70 (1H, d, J = 8.5), 7.75 (1H, dd, J = 7.0, 4.5). ^{13}C NMR (101 MHz, DMSO-d_6): 18.3, 19.3, 25.0, 25.5, 27.3, 27.6, 27.9, 28.0, 28.2, 28.4 (2C), 29.6, 35.1, 38.1, 40.0, 56.5, 58.6, 63.3, 68.9, 125.8, 128.0, 129.0, 139.8, 156.2, 170.6, 171.0. ESI MS: 526 ([M + Na]$^+$). HR ESI MS: calcd for $C_{28}H_{45}O_5N_3Na$, 526.32514; found, 526.32488.

*(4S,20S,21S)-20-Benzyl-21-hydroxy-4-isopropyl-1,3,6,19-tetraazacyclodocosane-2,5,18-trione (**34**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.73−0.87 (6H, m), 1.03−1.59 (19H, m), 1.73−1.93 (2H, m), 2.10−2.20 (1H, m), 2.61−2.84 (4H, m), 3.09−3.16 (1H, m), 3.56−3.62 (1H, m), 3.78 (1H, t, J = 8.1), 3.93 (1H, q, J = 8.3), 5.17 (1H, d, J = 5.8), 6.00−6.04 (1H, m), 6.19 (1H, d, J = 8.9), 7.10−7.35 (5H, m), 7.56−7.74 (2H, m). ^{13}C NMR (101 MHz, DMSO-d_6): 18.3, 19.5, 25.0, 26.0, 27.8 (2C), 28.0, 28.2, 28.5, 28.8, 28.9, 29.1, 30.0, 35.2, 37.3, 37.9, 40.0, 43.5, 52.9, 59.0, 70.7, 125.9, 128.1, 129.2, 139.4, 158.0, 172.0, 172.2. ESI MS: 525 ([M + Na]$^+$). HR ESI MS: calcd for $C_{28}H_{46}O_4N_4Na$, 525.34113; found, 525.34071.

*(12S,16S,17S)-17-Benzyl-16-hydroxy-12-isopropyl-1,4,7-trioxa-10,13,18-triazacyclohenicosane-11,14,19-trione (**35**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.82 (6H, d, J = 6.9), 2.04−2.14 (1H, m), 2.14−2.46 (5H, m), 2.63−2.74 (1H, m), 2.84 (1H, dd, J = 13.9, 4.9), 2.91−3.02 (1H, m), 3.36−3.48 (10H, m), 3.56−3.63 (3H, m), 3.93−4.01 (2H, m), 4.05 (1H, dd, J = 8.6, 5.7), 7.12−7.18 (1H, m), 7.18−7.28 (4H, m), 7.62 (1H, d, J = 8.7), 7.76 (1H, d, J = 9.2), 7.91 (1H, t, J = 5.3). ^{13}C NMR (101 MHz, DMSO-d_6): 17.6, 19.3, 29.4, 36.2, 36.4, 39.9, 53.3, 58.0, 66.9, 68.4, 68.9, 69.6, 69.7, 70.1, 70.4, 125.8, 128.0, 128.9, 139.5, 170.6, 170.7, 171.1. ESI MS: 516 ([M + Na]$^+$). HR ESI MS: calcd for $C_{25}H_{39}O_7N_3Na$, 516.26802; found, 516.26782.

*(3S,8R)-8-Benzyl-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**36**).* ^{1}H NMR (400 MHz, DMSO-d_6): 0.71−0.88 (6H, m), 0.95−1.53 (18H, m), 1.79−1.91 (1H, m), 1.92−2.04 (2H, m), 2.05−2.20 (3H, m), 2.57−2.77 (3H, m), 2.89−3.09 (1H, m), 3.78−3.90 (2H, m), 7.60 (1H, d, J = 7.9), 7.09−7.28 (5H, m), 7.67 (1H, dd, J = 8.5, 5.6), 7.91−7.97 (1H, m). ^{13}C NMR (101 MHz, DMSO-d_6): 18.0, 19.3, 24.8, 25.1, 26.4, 26.8, 27.1, 27.3, 27.4, 27.8, 28.0, 35.0, 38.0, 39.8, 49.2, 57.8, 58.8, 125.9, 128.0, 129.0, 139.1, 170.8, 172.0, 172.2. ESI MS: 494 ([M + Na]$^+$). HR ESI MS: calcd for $C_{28}H_{45}O_3N_3Na$, 494.33531; found, 494.33495.

Precursors of Macrocyclic Inhibitors. *Preparation of Statine Analogues. t-Butyl (S)-(1-(2,2-Dimethyl-4,6-dioxo-1,3-dioxan-5-yl)-3-(naphthalen-2-yl)-1-oxopropan-2-yl)carbamate (**38a**).* A solution of Boc-3-(2-naphthyl)-L-alanine (**37a**) (0.32 g, 1.00 mmol), Meldrum's acid (0.15 g, 1.05 mmol, 1.05 equiv), and DMAP (0.24 g, 2.00 mmol, 2.0 equiv) was prepared in anhydrous dichloromethane (DCM) (5 mL), and dicyclohexylcarbodiimide (DCC, 0.22 g, 1.05 mmol, 1.05 equiv) was added. The reaction mixture was stirred overnight and filtered, and the solid was washed with DCM (2 × 10 mL). The organic phase was washed with 10% $KHSO_4$ (2 × 20 mL), water (2 × 20 mL), and brine (20 mL). The solvent was evaporated to yield crude **38a**, which was used in the next step without further purification.

*t-Butyl (S)-3-Hydroxy-2-(naphthalen-2-ylmethyl)-5-oxo-2,5-dihydro-1H-pyrrole-1-carboxylate (**39a**).* The crude product **38a** (0.44 g, 1.00 mmol) was dissolved in EtOAc (10 mL), and the solution was heated to reflux for 30 min. After cooling, the product was extracted into the aqueous phase using 5% $NaHCO_3$ (3 × 10 mL). The aqueous phase was then acidified using $KHSO_4$ to pH 2.

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

The product was then extracted using EtOAc (3 × 15 mL). The extract was evaporated to dryness to yield crude **39a**.

*t-Butyl (2S,3S)-3-Hydroxy-2-(naphthalen-2-ylmethyl)-5-oxopyrrolidine-1-carboxylate (**40a**).* A solution of crude **39a** (0.34 g, 1.00 mmol) in a mixture of DCM (10 mL) and AcOH (1 mL) was cooled (0 °C), and $NaBH_4$ (0.08 g, 2.00 mmol, 2.0 equiv) was added portionwise. The mixture was stirred overnight at RT. It was then poured into cold water, and the organic layer was washed with water, dried with $MgSO_4$, and evaporated to dryness. Chromatography of the residue on silica gel (cyclohexane/EtOAc 2:1) afforded purified **40a** (0.20 g, 60% yield over 3 steps). ^{1}H NMR (400 MHz, DMSO-d_6): 1.46 (9H, s), 2.10 (1H, br s), 2.40−2.50 (1H, m), 2.55−2.63 (1H, m), 3.29 (2H, d, J = 6.0), 4.43−4.57 (2H, m), 7.40−7.50 (3H, m), 7.73−7.83 (4H, m). ^{13}C NMR (101 MHz, DMSO-d_6): 27.9, 34.2, 40.1, 62.6, 65.6, 83.4, 125.7, 126.2, 127.5, 127.6, 127.9, 128.2 (2C), 132.2, 133.5, 135.3, 149.8, 171.3.

*(3S,4S)-4-t-Butoxycarbonylamino-3-hydroxy-5-(2-napthyl)-pentanoic Acid (**41a**).* A solution of **40a** (0.20 g, 0.60 mmol) in acetone (5 mL) was treated dropwise with NaOH (1 M, 1.20 mL, 1.20 mmol, 2.0 equiv). After stirring for 30 min at RT, the solution was acidified to pH 3 with 1 M HCl and extracted with EtOAc (3 × 10 mL) to afford **41a** (0.20 g, 92% yield). ^{1}H NMR (400 MHz, DMSO-d_6): 1.26 (9H, s), 2.23 (1H, dd, J = 15.5, 9.2), 2.41 (1H, dd, J = 15.5, 3.6), 2.76 (1H, dd, J = 13.6, 9.4), 3.00 (1H, dd, J = 13.6, 5.1), 3.74−3.84 (1H, m), 3.90−3.98 (1H, m), 6.56 (1H, d, J = 9.3), 7.38−7.51 (3H, m), 7.71 (1H, s), 7.79−7.88 (3H, m). ^{13}C NMR (101 MHz, DMSO-d_6): 28.1, 36.1, 38.6, 55.7, 68.4, 77.6, 125.2, 125.8, 127.1, 127.2, 127.3, 127.4, 128.6, 131.9, 133.5, 135.3, 155.4, 173.1. EI MS: 359 ([M]$^+$). HR EI MS: calcd for $C_{20}H_{25}O_5N$, 359.1732; found, 359.1725. mp 180−181 °C. $[\alpha]_D^{20}$ (c 2.0, EtOH): −26.0°. EA: calcd for C, 66.83; H, 7.01; N, 3.90. Found C, 66.68; H, 7.08; N, 3.89.

*(3S,4S)-4-((t-Butoxycarbonyl)amino)-3-hydroxy-5-(thiazol-4-yl)-pentanoic Acid (**41b**).* ^{1}H NMR (400 MHz, DMSO-d_6): 1.32 (9H, s), 2.16 (1H, dd, J = 15.5, 9.3), 2.30 (1H, dd, J = 15.5, 3.2), 2.80 (1H, dd, J = 14.5, 9.1), 3.00 (1H, dd, J = 14.5, 4.8), 3.82−3.90 (2H, m), 6.42 (1H, d, J = 9.2), 7.28 (1H, d, J = 1.9), 9.00 (1H, d, J = 1.9). ^{13}C NMR (101 MHz, DMSO-d_6): 28.2, 32.1, 39.0, 53.8, 68.6, 77.6, 114.5, 153.1, 154.9, 155.3, 174.5. EI MS: 316 ([M]$^+$). HR EI MS: calcd for $C_{13}H_{20}O_5N_2S$, 316.1093; found, 316.1081. mp 145−148 °C. $[\alpha]_D^{20}$ (c 1.9, EtOH): −17.7°. EA: calcd for C, 49.35; H, 6.37; N, 8.85; S, 10.13. Found C, 48.96; H, 6.32; N, 8.48; S, 9.78.

*(3S,4R)-4-((t-Butoxycarbonyl)amino)-3-hydroxy-5-((4-methylbenzyl)thio)pentanoic Acid (**41c**).* ^{1}H NMR (400 MHz, DMSO-d_6): 1.41 (9H, s), 2.18 (1H, dd, J = 15.6, 8.9), 2.27 (3H, s), 2.31−2.41 (2H, m), 2.57 (1H, dd, J = 13.4, 6.0), 3.58−3.74 (3H, m), 3.99−4.07 (1H, m), 6.42 (1H, d, J = 9.3), 7.10 (2H, d, J = 7.7), 7.20 (1H, d, J = 8.0). ^{13}C NMR (101 MHz, DMSO-d_6): 20.7, 28.2, 31.9, 34.6, 38.6, 53.3, 67.7, 77.8, 128.8, 128.9, 135.4, 135.8, 155.6, 172.8. ESI MS: 392 ([M + Na]$^+$). HR ESI MS: calcd for $C_{18}H_{27}O_5NNaS$, 392.15021; found, 392.14989. mp 95−96 °C. $[\alpha]_D^{20}$ (c 8.8, MeOH): −23.1°. EA: calcd for C, 58.51; H, 7.37; N, 3.79; S, 8.68. Found C, 58.58; H, 7.44; N, 3.78; S, 8.55.

*(3S,4S)-4-t-Butoxycarbonylamino-3-hydroxy-5-(3,4-dichlorophenyl)-pentanoic Acid (**41d**).* ^{1}H NMR (400 MHz, DMSO-d_6): 1.29 (9H, s), 2.19 (1H, dd, J = 15.5, 9.1), 2.39 (1H, dd, J = 15.5, 3.6), 2.56 (1H, dd, J = 13.4, 10.1), 2.82 (1H, dd, J = 13.4, 4.2), 3.56−3.72 (1H, m), 3.83−3.96 (1H, m), 5.00 (1H, br s), 6.54 (1H, d, J = 9.4), 7.21 (1H, dd, J = 8.2, 1.9), 7.45−7.55 (2H, m), 11.20 (1H, br s). ^{13}C NMR (101 MHz, DMSO-d_6): 28.0, 35.1, 38.4, 55.6, 68.7, 77.7, 128.5, 129.7, 130.1, 130.5, 131.1, 140.9, 155.5, 173.0. CI MS: 400 ([M + Na]$^+$). HR EI MS: calcd for $C_{16}H_{21}O_5NCl_2$, 377.0795; found, 377.0796. mp 170−172 °C. $[\alpha]_D^{20}$ (c 1.2, EtOH): −25.0°. EA: calcd for C, 50.81; H, 5.60; N, 3.70; Cl, 18.75. Found C, 50.55; H, 5.60; N, 3.58; Cl, 18.43.

*(3S,4S)-4-t-Butoxycarbonylamino-3-hydroxy-5-(3,4-difluorophenyl)pentanoic Acid (**41e**).* ^{1}H NMR (400 MHz, DMSO-d_6): 1.29 (9H, s), 2.18 (1H, dd, J = 15.6, 9.1), 2.38 (1H, dd, J = 15.6, 3.7), 2.57 (1H, dd, J = 13.5, 10.2), 2.80 (1H, dd, J = 13.5, 4.5), 3.54−3.73 (1H, m), 3.81−3.97 (1H, m), 4.90 (1H, br s), 6.50 (1H, d, J = 9.4), 6.99−7.10 (1H, m), 7.17−7.40 (2H, m), 11.87 (1H, br s). ^{13}C NMR (101 MHz, DMSO-d_6): 28.0, 35.2, 38.5, 55.7, 68.6, 77.6, 116.9 (d, J_{CCCF} = 16.3), 117.4 (dd, J_{CCF} = 51.3, J_{CCCF} = 16.6), 125.9 (dd, J_{CCCF} = 6.2, J_{CCCCF} = 3.1), 137.4 (dd, J_{CCCF} = 5.8, J_{CCCCF} = 3.5), 148.4 (dd, J_{CF} = 243.6, J_{CCF} = 52.5), 148.6 (dd, J_{CF} = 243.8, J_{CCF} = 52.5), 155.5, 173.0. CI MS: 368 ([M + Na]$^+$). HR EI MS: calcd for $C_{16}H_{21}O_5NF_2$, 345.1315; found, 345.1387. mp 155−157 °C. $[\alpha]_D^{20}$ (c 1.33, EtOH): −26.4°. EA: calcd for C, 55.65; H, 6.13; N, 4.06; F, 11.00. Found C, 56.09; H, 6.25; N, 4.00; F, 10.79.

*(3S,4S)-4-t-Butoxycarbonylamino-3-hydroxy-5-(4-fluorophenyl)-pentanoic Acid (**41f**).* ^{1}H NMR (400 MHz, DMSO-d_6): 1.29 (9H, s), 2.18 (1H, dd, J = 15.5, 9.2), 2.37 (1H, dd, J = 15.5, 3.6), 2.55 (1H, dd, J = 13.6, 9.7), 2.80 (1H, dd, J = 13.6, 4.8), 3.54−3.71 (1H, m), 3.88 (1H, dddd, J = 8.8, 8.7, 3.2, 3.1), 6.48 (1H, d, J = 9.4), 7.01−7.15 (2H, m), 7.18−7.28 (2H, m). ^{13}C NMR (101 MHz, DMSO-d_6): 28.1, 35.1, 38.5, 55.8, 68.4, 77.6, 114.7 (d, J_{CCCF} = 20.9), 130.9 (d, J_{CCCF} = 7.8), 135.7 (d, J_{CCCCF} = 2.8), 155.5, 160.8 (d, J_{CF} = 241.4), 173.1. CI MS: 334 ([M + Li]$^+$). HR EI MS: calcd for $C_{16}H_{22}O_5NF$, 327.1482; found, 327.1468. mp 162−163 °C. $[\alpha]_D^{20}$ (c 2.5, EtOH): −31.2°. EA: calcd for C, 58.71; H, 6.77; N, 4.28; F, 5.80. Found C, 58.40; H, 6.71; N, 4.22; F, 5.67.

*Preparation of Aliphatic Aminoesters. t-Butyl 12-(2,2-Dimethyl-4,6-dioxo-1,3-dioxan-5-yl)-12-oxododecylcarbamate (**43**).* 12-(*t*-Butoxycarbonylamino)dodecanoic acid (**42**) (1.04 g, 3.29 mmol), dimethylaminopyridine (DMAP, 0.65 g, 5.28 mmol, 1.6 equiv), and Meldrum's acid (0.53 g, 3.62 mmol, 1.1 equiv) were partially dissolved in anhydrous DCM (30 mL) under an inert atmosphere. The mixture was cooled to 0 °C, and a solution of DCC (0.75 g, 3.62 mmol, 1.1 equiv) in DCM (20 mL) was added slowly to the suspension. The reaction was stirred at 0 °C under an inert atmosphere overnight. Dicyclohexylurea was filtered off and washed with DCM (2 × 10 mL). The solvent was evaporated, and the residue was chromatographed on silica gel ($CHCl_3$/acetone, gradient from 10:1 to 5:1) to afford desired derivative **43** (1.22 g, 84% yield) as a yellowish oil. ^{1}H NMR (400 MHz, $CDCl_3$): 1.19−1.36 (14H, m), 1.42 (9H, s), 1.50−1.58 (2H, m), 1.61 (3H, s), 1.65−1.70 (2H, m), 1.72 (3H, s), 1.84−1.92 (2H, m), 3.00−3.14 (2H, m), 3.38 (1H, s), 4.54 (1H, s). ^{13}C NMR (101 MHz, $CDCl_3$): 25.0, 25.7, 26.3, 26.9, 28.6, 29.4, 29.6, 29.8, 30.2, 31.1, 33.8, 35.9, 40.8, 42.3, 49.4, 79.2, 104.9, 157.2, 170.7, 198.5. ESI MS: 464 ([M + Na]$^+$). HR ESI MS: calcd for $C_{23}H_{39}O_7NNa$, 464.26187; found, 464.26224.

*t-Butyl 12-(2,2-Dimethyl-4,6-dioxo-1,3-dioxan-5-yl)-dodecylcarbamate (**44**).* The intermediate **43** (1.22 g, 2.72 mmol) was dissolved in DCM (30 mL) and cooled to −5 °C. Acetic acid (0.96 mL, 16.58 mmol, 6.0 equiv) was added dropwise, and sodium borohydride (0.31 g, 8.15 mmol, 3.0 equiv) was added in three portions. The mixture was stirred at −5 °C overnight and then washed with a saturated solution of NaCl (2 × 20 mL). The organic phase was dried over $MgSO_4$, filtered, and evaporated in vacuo. Column chromatography ($CHCl_3$/acetone 25:1) afforded desired derivative **44** (0.65 g, 56% yield) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$): 1.22−1.37 (16H, m), 1.38−1.50 (4H, m), 1.43 (9H, s), 1.75 (3H, s), 1.78 (3H, s), 2.05−2.14 (2H, m), 3.09 (2H, t, J = 7.0), 3.49 (1H, dd, J = 6.8, 3.3), 4.48 (1H, s). ^{13}C NMR (101 MHz, $CDCl_3$): 26.5, 26.7, 26.8, 27.0, 28.4, 29.2, 29.3, 29.5 (5C), 30.1, 40.7, 46.1, 79.0, 104.8, 156.0, 165.7. ESI MS: 450 ([M + Na]$^+$). HR ESI MS: calcd for $C_{23}H_{41}O_6NNa$, 450.28261; found, 450.28273.

*14-(t-Butoxycarbonylamino)tetradecanoic Acid (**45**).* Intermediate **44** (0.55 g, 1.28 mmol) was heated to 110 °C for 3 h in DMF/H_2O (10:1, 5 mL). After cooling, solvents were evaporated to obtain **45** (0.42 g, 94% yield) as a white solid. ^{1}H NMR (400 MHz, CD_3OD): 1.30 (18H, s), 1.43 (9H, s), 1.45−1.50 (2H, m), 1.55−1 65 (2H, m), 2.27 (2H, t, J = 7.4), 2.62 (1H, s), 3.01 (2H, t, J = 7.0). ^{13}C NMR (101 MHz, CD_3OD): 26.2, 27.9, 28.8, 30.3, 30.4, 30.6, 30.7 (6C), 31.0, 35.1, 41.4, 79.7, 158.6, 177.8. ESI MS: 342 ([M − H]$^-$). HR ESI MS: calcd for $C_{19}H_{36}O_4N$, 342.26498; found, 342.26445.

*Methyl 14-Aminotetradecanoate (**46**).* A solution of **45** (0.36 g, 1.04 mmol) in MeOH (5 mL) was added slowly to a cooled (0 °C) solution of thionyl chloride (0.31 mL, 4.17 mmol, 4.0 equiv) in MeOH (10 mL). The reaction was stirred at 0 °C for 1 h, and the organic solvent was evaporated. The residue was dissolved in

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

trifluoroacetic acid (TFA, 2 mL) and stirred at room temperature for 15 min. TFA was removed by flow of nitrogen. The residue was partially dissolved in a saturated solution of $NaHCO_3$ (10 mL) and extracted with EtOAc (3 × 15 mL). The organic phases were combined and dried over $MgSO_4$. After filtration, the solvent was evaporated to obtain **46** (0.17 g, 62% yield) as a white solid. ^{1}H NMR (400 MHz, CD_3OD): 1.25−1.39 (18H, m), 1.50−1.64 (4H, m), 2.31 (2H, t, J = 7.4), 3.75 (2H, t, J = 7.4), 3.65 (3H, s). ^{13}C NMR (101 MHz, CD_3OD): 26.0, 27.8, 30.2, 30.4, 30.5, 30.6 (2C), 30.7 (3C), 31.4, 34.8, 41.7, 51.9, 176.0. ESI MS: 258 ([M + H]$^+$). HR ESI MS: calcd for $C_{15}H_{31}O_2NH$, 258.42276; found, 258.42278.

*Procedures for Preparation of Macrocyclic Inhibitors with an Aliphatic Chain (**6**−**32**, **34** and **36**). Methyl (S)-12-(2-((t-Butoxycarbonyl)amino)-3-methylbutanamido) Dodecanoate (**49a**).* Boc-L-Val-OH (**47a**) (0.84 g, 3.88 mmol) and HBTU (1.48 g, 3.88 mmol, 1.0 equiv) were dissolved in anhydrous DMF (15 mL), and DIEA (1.97 mL, 11.32 mmol, 3.0 equiv) was added. The reaction mixture was stirred at room temperature for 25 min, and a solution of methyl 12-aminododecanoate (**48**) (0.86 g, 3.23 mmol, 0.8 equiv) in anhydrous DMF (5 mL) was added. The reaction mixture was stirred at room temperature for 6 h under an inert atmosphere. The reaction mixture was diluted with $CHCl_3$ (30 mL) and washed with saturated $NaHCO_3$ (2 × 25 mL), 10% $KHSO_4$ (2 × 25 mL), and brine (35 mL). The mixture was dried over $MgSO_4$, and the organic solvent was evaporated in vacuo. Column chromatography on silica gel ($CHCl_3$/MeOH 30:1) afforded the desired product **49a** (1.23 g, 89% yield) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$): 0.90−0.95 (6H, m), 1.20−1.34 (14H, m), 1.42 (9H, s), 1.45−1.51 (2H, m), 1.56−1.64 (2H, m), 2.03−2.14 (1H, m), 2.29 (2H, t, J = 7.7), 3.14−3.31 (2H, m), 3.66 (3H, s), 3.86 (1H, dd, J = 8.8, 6.6), 5.20 (1H, s), 6.31 (1H, s). ^{13}C NMR (101 MHz, $CDCl_3$): 18.1, 19.4, 25.0, 27.0, 28.4, 29.2, 29.3 (3C), 29.5, 29.6, 29.7, 31.0, 34.2, 39.5, 51.5, 60.2, 156.0, 171.7, 174.4. ESI MS: 451 ([M + Na]$^+$).

*Methyl (S)-12-(2-((t-Butoxycarbonyl)amino)-3-(pyridin-3-yl)-propanamido) Dodecanoate (**49b**).* Boc-L-3Pal-OH (**47b**) (0.18 g, 0.68 mmol) and HBTU (0.26 g, 0.68 mmol, 1.0 equiv) were dissolved in anhydrous DMF (3 mL), and DIEA (0.34 mL, 1.98 mmol, 3.0 equiv) was added. The reaction mixture was stirred at room temperature for 25 min, and a solution of methyl 12-amino-dodecanoate (**48**) (0.15 g, 0.56 mmol, 0.8 equiv) in anhydrous DMF (2 mL) was added. The reaction mixture was stirred at room temperature for 6 h under an inert atmosphere. The reaction mixture was diluted with $CHCl_3$ (15 mL) and washed with saturated $NaHCO_3$ (2 × 15 mL), 10% $KHSO_4$ (2 × 15 mL), and brine (15 mL). The mixture was dried over $MgSO_4$, and the organic solvent was evaporated in vacuo. Column chromatography ($CHCl_3$/MeOH 25:1) afforded the desired product **49b** (0.23 g, 85% yield) as an amorphous solid. ^{1}H NMR (400 MHz, $CDCl_3$): 1.12−1.29 (14H, m), 1.34 (9H, s), 1.35−1.37 (2H, m), 1.58 (2H, p, J = 7.5), 2.27 (2H, t, J = 7.5), 2.91−2.98 (1H, m), 3.05−3.21 (3H, m), 3.63 (3H, s), 4.29−4.36 (1H, m), 5.38 (1H, d, J = 8.6), 6.46 (1H, s), 7.20 (1H, dd J = 7.9, 4.8), 7.54−7.57 (1H, m), 8.41−8.45 (2H, m). ^{13}C NMR (101 MHz, $CDCl_3$): 25.0, 26.9, 28.3, 29.2, 29.3 (2C), 29.5 (2C), 29.6, 29.8, 34.2, 36.1, 39.7, 51.5, 55.5, 80.3, 123.6, 133.1, 137.5, 147.8, 150.2, 155.5, 170.7, 174.4. ESI MS: 478 ([M + H]$^+$). HR ESI MS: calcd for $C_{26}H_{44}O_5N_3$, 478.32755; found, 478.32760.

*Methyl (6S,7S,11S)-6-Benzyl-7-hydroxy-2,2-dimethyl-4,9,12-trioxo-11-isopropyl-3-oxa-5,10,13-triazapentacosan-25-oate (**50a**).* Compound **49a** (0.18 g, 0.42 mmol) was dissolved in DCM (2 mL), and TFA (0.2 mL) was added. The reaction mixture was stirred at room temperature for 1 h. The reaction mixture was diluted with CH_2Cl_2 (10 mL), washed with saturated $NaHCO_3$ (2 × 10 mL), and dried over $MgSO_4$. The organic solvent was evaporated in vacuo to obtain crude amine. Boc-phenylstatine-OH (**52**) (0.16 g, 0.51 mmol) and HBTU (0.2 g, 0.51 mmol, 1.0 equiv) were dissolved in anhydrous DMF (4 mL), and triethylamine (0.19 mL, 1.36 mmol, 2.5 equiv) was added. The reaction mixture was stirred at room temperature for 30 min, and then a solution of the amine (0.14 g, 0.42 mmol, 0.8 equiv) in anhydrous DMF (2 mL) was added. The reaction mixture was stirred at room temperature for 14 h under an inert atmosphere. The organic solvent was evaporated in vacuo. The residue was chromatographed on silica gel ($CHCl_3$/MeOH 10:1) to obtain the desired product (0.16 g, 61% yield) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$): 0.87−0.94 (6H, m), 1.21−1.33 (14H, m), 1.37 (9H, s), 1.42−1.49 (2H, m), 1.56−1.64 (2H, m), 2.01−2.14 (1H, m), 2.25−2.33 (3H, m), 2.50 (1H, dd, J = 15.0, 9.2), 2.80 (1H, s), 2.89 (2H, d, J = 7.5), 3.09−3.27 (2H, m), 3.65 (3H, s), 3.68−3.74 (1H, m), 3.94−4.01 (1H, m), 4.17 (1H, dd, J = 8.4, 7.0), 5.05 (1H, d, J = 9.7), 6.46 (1H, s), 6.77 (1H, d, J = 8.4), 7.14−7.29 (5H, m). ^{13}C NMR (101 MHz, $CDCl_3$): 18.3, 19.4, 25.1, 27.0, 29.3 (2C), 29.4, 29.5 (2C), 29.6 (2C), 31.0, 34.2, 38.7, 39.8, 40.5, 51.6, 55.6, 59.0, 68.7, 79.7, 126.5, 128.5, 129.4, 138.3, 156.4, 171.0, 172.6, 174.5. ESI MS: 642 ([M + Na]$^+$).

*Methyl (6S,7S,11S)-6-Benzyl-7-hydroxy-2,2-dimethyl-4,9,12-trioxo-11-(pyridin-3-ylmethyl)-3-oxa-5,10,13-triazapentacosan-25-oate (**50b**).* Compound **49b** (0.18 g, 0.38 mmol) was dissolved in DCM (2 mL), and TFA (0.2 mL) was added. The reaction mixture was stirred at room temperature for 1 h. The reaction mixture was diluted with CH_2Cl_2 (10 mL), washed with saturated $NaHCO_3$ (2 × 10 mL), and dried over $MgSO_4$. The organic solvent was evaporated in vacuo to obtain crude amine. Boc-phenylstatine-OH (**52**) (0.14 g, 0.46 mmol) and HBTU (0.17 g, 0.46 mmol, 1.0 equiv) were dissolved in anhydrous DMF (3 mL), and triethylamine (0.17 mL, 1.23 mmol, 2.7 equiv) was added. The reaction mixture was stirred at room temperature for 30 min, and a solution of the amine (0.15 g, 0.38 mmol, 0.8 equiv) in anhydrous DMF (2 mL) was added. The reaction mixture was stirred at room temperature for 14 h under an inert atmosphere. The organic solvent was evaporated in vacuo. The residue was chromatographed on silica gel ($CHCl_3$/MeOH 15:1) to obtain the desired product (0.13 g, 51% yield) as an amorphous solid. ^{1}H NMR (400 MHz, $CDCl_3$): 1.11−1.36 (25H, m), 1.55−1.62 (2H, m), 2.23−2.28 (1H, m), 2.27 (2H, t, J = 7.5), 2.42 (1H, dd, J = 14.9, 9.0), 2.82−2.88 (2H, m), 2.96−3.15 (4H, m), 3.64 (3H, s), 3.67−3.74 (1H, m), 3.92−3.97 (1H, m), 4.77 (1H, dd, J = 7.5, 7.2), 4.84−4.89 (1H, m), 5.11 (1H, d, J = 9.7), 5.79−5.81 (1H, m), 7.03−7.08 (1H, m), 7.11−7.24 (6H, m), 7.51−7.55 (1H, m), 8.36−8.40 (2H, m). ^{13}C NMR (101 MHz, $CDCl_3$): 25.0, 26.9, 28.4, 29.2, 29.3 (2C), 29.5 (2C), 29.6 (2C), 34.2, 36.0, 38.7, 39.7, 40.9, 51.5, 54.0, 55.6, 68.5, 79.6, 123.6, 126.4, 128.5, 129.4, 132.8, 137.6, 138.3, 147.8, 150.2, 156.3, 170.2, 172.1, 174.4. ESI MS: 691 ([M + Na]$^+$). HR ESI MS: calcd for $C_{37}H_{56}O_7N_4Na$, 691.40412; found, 691.40431.

*(6S,7S,11S)-6-Benzyl-7-hydroxy-11-isopropyl-2,2-dimethyl-4,9,12-trioxo-3-oxa-5,10,13-triazapentacosan-25-oic Acid (**51a**).* To a solution of methylester **50a** (0.15 g, 0.24 mmol) in THF (2 mL), a solution of 1 M LiOH (2 mL, 2.0 mmol, 8.3 equiv) was added dropwise. The reaction mixture was stirred at room temperature for 2 h and then acidified with 1 M hydrochloric acid (5 mL). $CHCl_3$ (10 mL) was added. The layers were separated, and the product was extracted with $CHCl_3$ (3 × 10 mL). The organic layers were combined and dried over $MgSO_4$, and the solvent was removed in vacuo. The residue was chromatographed on silica gel ($CHCl_3$/MeOH 20:1) to yield the desired acid **51a** (0.13 g, 89% yield) as white crystals. ^{1}H NMR (400 MHz, $CDCl_3$): 0.87−0.94 (6H, m), 1.21−1.35 (16H, m), 1.38 (9H, s), 1.41−1.50 (2H, m), 1.57−1.66 (2H, m), 2.04 (1H, s), 2.29−2.36 (3H, m), 2.47−2.57 (1H, m), 2.89 (2H, d, J = 7.4), 3.10−3.30 (2H, m), 3.48 (1H, s), 3.67−3.76 (1H, m), 3.94−4.02 (1H, m), 4.18−4.24 (1H, m), 5.01−5.12 (1H, m), 6.73 (1H, m), 7.16−7.29 (5H, m). ^{13}C NMR (101 MHz, $CDCl_3$): 18.5, 19.4, 24.8, 26.8, 28.5, 28.9, 29.0 (2C), 29.1, 29.2, 29.3, 29.8, 31.0, 34.1, 38.7, 39.8, 40.2, 55.7, 59.2, 68.6, 79.7, 126.5, 128.3, 129.5, 138.3, 156.4, 171.5, 173.1, 175.6, 178.0. ESI MS: 628 ([M + Na]$^+$).

*(6S,7S,11S)-6-Benzyl-7-hydroxy-2,2-dimethyl-4,9,12-trioxo-11-(pyridin-3-ylmethyl)-3-oxa-5,10,13-triazapentacosan-25-oic Acid (**51b**).* To a solution of methylester **50b** (0.13 g, 0.19 mmol) in THF (2 mL), a solution of 1 M LiOH (2 mL, 2.0 mmol, 10.5 equiv) was added dropwise. The reaction mixture was stirred at room temperature for 2 h and then acidified with 1 M hydrochloric acid (5 mL). $CHCl_3$ (10 mL) was added. The layers were separated, and the product was extracted with $CHCl_3$ (3 × 10 mL). The organic layers were combined and dried over $MgSO_4$, and the solvent was removed

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

in vacuo. The residue was chromatographed on silica gel ($CHCl_3$/MeOH gradient from 15:1 to 5:1) to yield the desired acid **51b** (0.10 g, 77% yield) as an amorphous solid. ^{1}H NMR (400 MHz, $CDCl_3$): 1.12−1.35 (2SH, m), 1.54−1.62 (2H, m), 2.21−2.27 (1H, m), 2.28 (2H, t, J = 7.5), 2.44 (1H, dd, J = 14.9, 9.0), 2.83−2.88 (2H, m), 2.96−3.14 (4H, m), 3.68−3.75 (1H, m), 3.92−3.98 (1H, m), 4.78 (1H, dd, J = 7.5, 7.3), 4.84−4.88 (1H, m), 5.11 (1H, d, J = 9.7), 5.77−5.81 (1H, m), 7.02−7.09 (1H, m), 7.12−7.24 (6H, m), 7.51−7.54 (1H, m), 8.36−8.41 (2H, m). ^{13}C NMR (101 MHz, $CDCl_3$): 25.0, 26.9, 28.4, 29.2, 29.3 (2C), 29.5 (2C), 29.6 (2C), 34.2, 36.1, 38.7, 39.7, 41.0, 54.0, 55.7, 68.5, 79.6, 123.6, 126.4, 128.5, 129.3, 132.8, 137.6, 138.3, 147.8, 150.2, 156.4, 170.3, 172.1, 172.2. ESI MS: 677 ([M + Na]$^+$). HR ESI MS: calcd for $C_{36}H_{54}O_7N_4Na$, 677.38847; found, 677.38849.

*(3S,7S,8S)-8-Benzyl-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (**9**).* To a solution of acid **51a** (92 mg; 0.14 mmol) and TSTU (42 mg, 0.14 mmol, 1.0 equiv) in anhydrous DMF (2 mL), DIEA (26 μL, 0.14 mmol, 1.0 equiv) was added. The reaction mixture was stirred at room temperature for 2 h under an inert atmosphere. The reaction mixture was diluted with $CHCl_3$ (15 mL), washed with brine (2 × 10 mL), and dried over $MgSO_4$. The organic solvent was evaporated in vacuo. The residue was dissolved in TFA (1 mL), and the reaction mixture was stirred at room temperature for 15 min. TFA was removed by flow of nitrogen to obtain crude amine. The crude amine was dissolved in anhydrous DMF (5 mL) and added dropwise overnight to a solution of DIEA (5 mL) in anhydrous DMF (5 mL) under stirring. The solvent was evaporated, and the crude product was purified by preparative HPLC to afford **9** (29 mg, 41% yield). Spectral data for **9** are reported in the part characterization of macrocyclic inhibitors.

*(3S,7S,8S)-8-Benzyl-7-hydroxy-3-(pyridin-3-ylmethyl)-1,4,9-triazacyclohenicosane-2,5,10-trione (**25**).* To a solution of acid **51b** (90 mg, 0.14 mmol) and TSTU (42 mg, 0.14 mmol, 1.0 equiv) in anhydrous DMF (2 mL), DIEA (24 μL, 0.14 mmol, 1.0 equiv) was added. The reaction mixture was stirred at room temperature for 2 h under an inert atmosphere. The reaction mixture was diluted with $CHCl_3$ (15 mL), washed with brine (2 × 10 mL), and dried over $MgSO_4$. The organic solvent was evaporated in vacuo. The residue was dissolved in TFA (1 mL), and the mixture was stirred at room temperature for 15 min. TFA was removed by flow of nitrogen to obtain crude amine. The crude amine was dissolved in anhydrous DMF (5 mL) and added dropwise overnight to a solution of DIEA (5 mL) in anhydrous DMF (5 mL) under stirring. The solvent was evaporated, and crude product was purified by preparative HPLC to afford **25** (35 mg, 47% yield). Spectral data for **25** are reported in the part characterization of macrocyclic inhibitors.

*Procedures for Preparation of Macrocycle with a PEG Chain (**35**). (S)-Methyl 2-((3S,4S)-4-(t-Butoxycarbonylamino)-3-hydroxy-5-phenylpentanamido)-3-methylbutanoate (**53**).* Boc-phenylstatine-OH (**52**) (0.19 g, 0.61 mmol) and HBTU (0.23 g, 0.61 mmol, 1.0 equiv) were dissolved in anhydrous DMF (2.5 mL), and DIEA (0.36 mL, 2.05 mmol, 3.4 equiv) was added. The reaction mixture was stirred at room temperature for 30 min, and a solution of H-Val-OMe (90 mg, 0.50 mmol, 0.8 equiv) in anhydrous DMF (2 mL) was added. The reaction mixture was stirred at room temperature for 12 h under an inert atmosphere. The reaction mixture was diluted with $CHCl_3$ (30 mL) and washed with saturated $NaHCO_3$ (2 × 20 mL), 10% $KHSO_4$ (2 × 20 mL), and brine (30 mL). The mixture was dried over $MgSO_4$ and the organic solvent was evaporated in vacuo. The residue was chromatographed on silica gel ($CHCl_3$/MeOH 20:1) to obtain the desired product **53** (0.16 g, 62% yield). ^{1}H NMR (400 MHz, $CDCl_3$): 0.90 (3H, d, J = 6.9), 0.93 (3H, d, J = 6.9), 1.40 (9H, s), 2.08−2.20 (1H, m), 2.29 (1H, dd, J = 15.2, 3.1), 2.52 (1H, dd, J = 15.2, 9.6), 2.85 (1H, s), 2.91 (2H, d, J = 7.6), 3.72 (3H, s), 3.73−3.81 (1H, m), 3.96 (1H, d, J = 8.1), 4.49 (1H, dd, J = 8.6, 4.9), 4.99 (1H, d, J = 9.4), 6.53 (1H, d, J = 8.4), 7.16−7.31 (5H, m). ^{13}C NMR (101 MHz, $CDCl_3$): 17.9, 19.1, 28.5, 31.2, 38.7, 40.0, 52.4, 55.6, 57.3, 68.2, 79.7, 126.5, 128.6, 129.5, 138.3, 156.4 172.3, 172.9. ESI MS: 445 ([M + Na]$^+$). HR ESI MS: calcd for $C_{22}H_{34}O_6N_2Na$, 445.23091, found, 445.23092.

*(19S,20S,24S)-Methyl 19-Benzyl-20-hydroxy-24-isopropyl-2,2-dimethyl-4,17,22-trioxo-3,8,11,14-tetraoxa-5,18,23-triazapentacosan-25-oate (**54**).* Compound **53** (0.15 g, 0.33 mmol) was dissolved in TFA (1 mL), and the reaction mixture was stirred at room temperature for 15 min. TFA was removed by flow of nitrogen to obtain crude amine. BocHN-PEG-3-CH_2CH_2COOH (0.10 g, 0.33 mmol) and HBTU (0.12 g, 0.33 mmol, 1.0 equiv) were dissolved in anhydrous DMF (2.5 mL), and DIEA (0.17 mL, 0.99 mmol, 3.0 equiv) was added. The reaction mixture was stirred at room temperature for 30 min, and a solution of the amine (0.10 g, 0.33 mmol, 1.0 equiv) in anhydrous DMF (2 mL) was added. The reaction mixture was stirred at room temperature for 16 h under an inert atmosphere. Column chromatography on silica gel ($CHCl_3$/MeOH 20:1) afforded the desired product **54** (0.15 g, 79% yield). ^{1}H NMR (400 MHz, $CDCl_3$): 0.92 (3H, d, J = 6.9), 0.94 (3H, d, J = 6.9), 1.43 (9H, s), 2.10−2.21 (1H, m), 2.38 (1H, dd, J = 14.9, 4.7), 2.47−2.51 (2H, m), 2.53 (1H, dd, J = 15.0, 8.9), 2.93−2.98 (4H, m), 3.29 (2H, t, J = 5.1), 3.53 (2H, t, J = 5.2), 3.58−3.65 (8H, m), 3.70 (3H, s), 3.99−4.04 (1H, m), 4.15 (1H, dd, J = 16.1, 8.4), 4.44 (1H, dd, J = 8.4, 5.2), 7.23−7.28 (5H, m). ^{13}C NMR (101 MHz, $CDCl_3$): 18.0, 19.1, 28.6, 30.8, 36.8, 37.8, 40.4, 52.3, 54.8, 57.7, 67.3, 69.0, 70.3, 70.4 (4C), 70.5, 79.6, 126.6, 128.6, 129.4, 138.2, 156.3, 172.3, 172.9, 173.3. ESI MS: 648 ([M + Na]$^+$). HR ESI MS: calcd for $C_{31}H_{51}O_{10}N_3Na$, 648.34667; found, 648.34678.

*(12S,16S,17S)-17-Benzyl-16-hydroxy-12-isopropyl-1,4,7-trioxa-10,13,18-triazacyclo Henicosane-11,14,19-trione (**35**).* To a solution of corresponding methylester **54** (65 mg, 0.10 mmol) in THF (2 mL), a solution of LiOH (7 mg, 0.15 mmol, 1.5 equiv) in water (2 mL) was added dropwise. The reaction mixture was stirred at room temperature for 2 h, and then the organic solvent was evaporated in vacuo. The residue was dissolved in TFA (2 mL), and the reaction mixture was stirred at room temperature for 15 min. TFA was removed by flow of nitrogen to obtain crude amino acid salt. A solution of DIEA (0.37 mL, 2.00 mmol, 20 equiv) and DPPA (30 mg, 0.11 mmol, 1.1 equiv) in anhydrous DMF (10 mL) was slowly added to a solution of the crude amino acid salt in DMF (5 mL). The reaction mixture was stirred overnight. The solvent was evaporated and crude product was purified by preparative HPLC to afford desired compound **35** (27 mg, 55% yield). Spectral data for **35** are reported in the part characterization of macrocyclic inhibitors.

*Procedures for Preparation of Macrocycle with a Carbamate Moiety (**33**). t-Butyl (S)-(1-((11-Hydroxyundecyl)amino)-3-methyl-1-oxobutan-2-yl)carbamate (**57**).* Boc-L-Val-OSu (**55**) (1.00 g, 3.18 mmol) and 11-aminoundecan-1-ol (**56**) (0.65 g, 3.50 mmol, 1.1 equiv) were dissolved in anhydrous DMF (10 mL), and DIEA (1.19 mL, 6.40 mmol, 2.0 equiv) was added. The reaction mixture was stirred at room temperature overnight. The reaction mixture was diluted with EtOAc (50 mL) and washed with saturated $NaHCO_3$ (3 × 50 mL), 10% $KHSO_4$ (3 × 50 mL), and brine (35 mL). The mixture was dried over $MgSO_4$ and the organic solvent was evaporated in vacuo to yield the desired product **57** (1.13 g, 93% yield) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$): 0.90 (3H, d, J = 6.8), 0.93 (3H, d, J = 6.8), 1.20−1.34 (12H, m), 1.42 (9H, s), 1.44−1.58 (4H, m), 1.78−1.94 (1H, m), 2.00−2.15 (1H, m), 3.13−3.30 (2H, m), 3.61 (2H, t, J = 6.6), 3.82 (1H, dd, J = 9.0, 6.6), 5.16 (1H, d, J = 9.1), 6.21 (1H, t, J = 5.8). ^{13}C NMR (101 MHz, $CDCl_3$): 18.1, 19.4, 25.8, 26.9, 28.4, 29.3, 29.5 (3C), 29.6, 30.9, 32.9, 39.6, 60.3, 63.0, 79.9, 156.1, 171.7. ESI MS: 409 ([M + Na]$^+$). HR ESI MS: calcd for $C_{21}H_{42}O_4N_2Na$, 409.30368; found, 409.30348.

*t-Butyl ((2S,3S)-3-Hydroxy-5-(((S)-1-((11-hydroxyundecyl)-amino)-3-methyl-1-oxobutan-2-yl)amino)-5-oxo-1-phenylpentan-2-yl)carbamate (**58**).* Compound **57** (1.0 g, 2.59 mmol) was dissolved in DCM (5 mL), and TFA (5 mL) was added. The reaction mixture was stirred at room temperature for 1 h. The organic solvent was evaporated in vacuo to obtain crude amine. Boc-phenylstatine-OH (**52**) (0.72 g, 2.33 mmol) and TSTU (0.70 g, 2.33 mmol, 1.0 equiv) were dissolved in anhydrous DMF (5 mL), and DIEA (1.44 mL, 7.77 mmol, 3.0 equiv) was added. The reaction mixture was stirred at room temperature for 1 h. A solution of the amine in anhydrous DMF (2 mL) was added. The reaction mixture

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

was stirred at room temperature for 2 h. The reaction mixture was then diluted with EtOAc (50 mL) and washed with saturated $NaHCO_3$ (3 × 50 mL), 10% $KHSO_4$ (3 × 50 mL), and brine (35 mL). The mixture was dried over $MgSO_4$ and the organic solvent was evaporated in vacuo. Column chromatography on silica gel (DCM/MeOH 20:1) afforded the desired product **58** (1.0 g, 75% yield). ^{1}H NMR (400 MHz, CD_3OD): 0.94 (3H, d, J = 6.5), 0.96 (3H, d, J = 6.5), 1.21−1.33 (14H, m), 1.35 (9H, s), 1.41−1.57 (5H, m), 2.12 (1H, dq, J = 13.5, 6.8), 2.45 (1H, dd, J = 7.0, 3.1), 2.77 (1H, dd, J = 13.7, 9.5), 2.87 (1H, dd, J = 13.7, 5.4), 3.15 (2H, ddp, J = 20.3, 13.4, 7.1), 3.53 (2H, t, J = 6.7), 3.74 (1H, ddd, J = 9.6, 5.4, 2.1), 4.06 (1H, td, J = 7.0, 2.0), 4.11 (1H, d, J = 6.0), 7.11−7.28 (5H, m). ^{13}C NMR (101 MHz, CD_3OD): 18.2, 19.7, 27.0, 28.0, 28.8, 30.4 (2C), 30.6, 30.7 (2C), 31.4, 33.7, 39.2, 40.4, 41.5, 57.0, 60.7, 63.0, 70.8, 80.1, 127.2, 129.2, 130.4, 140.2, 158.4, 173.6, 174.0. ESI MS: 600 ([M + Na]$^+$). HR ESI MS: calcd for $C_{32}H_{55}O_6N_3Na$, 600.39831; found, 600.39754.

*(4S,5S,9S)-4-Benzyl-5-hydroxy-9-isopropyl-1-oxa-3,8,11-triazacyclodocosane-2,7,10-trione (**33**).* Compound **58** (50 mg, 0.09 mmol) and disuccinimidyl carbonate (33 mg, 0.13 mmol, 1.5 equiv) were dissolved in anhydrous DMF (2 mL). The reaction mixture was stirred at room temperature overnight. The reaction mixture was then diluted with EtOAc (25 mL) and washed with saturated $NaHCO_3$ (3 × 25 mL), 10% $KHSO_4$ (3 × 25 mL), and brine (20 mL). It was then dried over $MgSO_4$, and the organic solvent was evaporated in vacuo. The intermediate was then dissolved in DCM (2 mL), and TFA (2 mL) was added. The reaction mixture was stirred at room temperature for 20 min. The organic solvent was evaporated in vacuo to obtain crude amine. The crude amine was dissolved in anhydrous DMF (2 mL) and added dropwise overnight to a solution of DIEA (1 mL) in anhydrous DMF (5 mL) under stirring. The solvent was evaporated, and the crude product was purified by preparative HPLC to afford **33** (23 mg, 52% yield). Spectral data for **33** are reported in the part characterization of macrocyclic inhibitors.

Enzymatic Assays. The enzymatic activity of human CatD was measured in a kinetic assay using the FRET substrate Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu.[38,40] The assay was performed in 96-well format with a total reaction volume of 100 μL at 37 °C. The kinetics of the product release were continuously monitored in an Infinite M1000 microplate reader (Tecan) at excitation and emission wavelengths of 330 and 420 nm, respectively. For inhibition measurements, 1 nM CatD was preincubated for 10 min with 0−200 μM inhibitors in 0.1 M sodium acetate, pH 4.2, 0.1% (w/v) PEG 1500, followed by addition of 25 μM substrate. The final concentration of DMSO in the assay did not exceed 2% (v/v). IC_{50} values were determined from residual velocities using dose−response plots; nonlinear regression was fitted using GraFit software (Erithacus). The mode of inhibition was analyzed at varying substrate concentrations, and the initial velocities of product release were interpreted using double reciprocal plots.

Screening of macrocyclic inhibitors against a panel of aspartic proteases was performed using a modified human CatD assay in 384-well format with a total reaction volume of 20 μL at 37 °C. Inhibitors were applied at a concentration of 50 nM. The assay conditions for individual proteases were as follows: 1.2 nM tick parasite CatD and 25 μM Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu substrate in 0.1 M sodium acetate, pH 4.0; 2.7 nM yeast proteinase A and 25 μM Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu substrate in 0.1 M sodium acetate, pH 5.0; 10 nM porcine pepsin and 10 μM Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu substrate in 0.1 M sodium acetate, pH 4.0; 30 nM human BACE1 and 25 μM Mca-Ser-Glu-Val-Asn-Leu-Asp-Ala-Glu-Phe-Arg-Lys(Dnp)-Arg-Arg substrate in 0.1 M sodium acetate, pH 4.6; 14 nM HIV-1 protease and 40 μM Abz-Tyr-Ile-Nle-Nph-Gln-Arg substrate in 0.1 M sodium acetate, pH 5.0, 1 mM EDTA; and 20 nM human renin and 10 μM Arg-Glu(EDANS)-Ile-His-Pro-Phe-His-Leu-Val-Ile-His-Thr-Lys(Dabcyl)-Arg substrate in 0.1 M MES, pH 6.0, 0.15 M NaCl.[41−43]

Screening of macrocyclic inhibitors against a panel of non-aspartic proteases was performed in 98-well format with a total reaction volume of 100 μL at 37 °C. Inhibitors were applied at a concentration of 1 μM. The assay conditions for individual proteases were as follows: 2 nM trypsin and 10 μM Z-Phe-Arg-AMC substrate in 0.1 M Tris-HCl, pH 7.5, 10 mM $CaCl_2$; 0.34 nM chymotrypsin and 10 μM Z-Phe-Arg-AMC substrate in 0.1 M Tris-HCl, pH 7.5; 1.2 nM papain and 10 μM Z-Phe-Arg-AMC substrate in 0.1 M sodium acetate, pH 5.5; 4.3 nM thermolysin, and 10 μM Abz-Ala-Gly-Leu-Ala-nitrobenzylamide in 0.1 M Tris-HCl, pH 7.5, 10 mM $CaCl_2$.[44,45]

All enzymatic activity measurements were performed in triplicate in at least two independent experiments.

Resistance of macrocycles to proteolysis was investigated using a previously described assay.[46] The reaction mixture, containing 25 μg macrocycle and 1 μg protease (trypsin, chymotrypsin, papain, thermolysin, or human CatD), was incubated in the corresponding protease activity assay buffer in a total volume of 150 μL at 26 °C for 18 h. The reaction was stopped by adding 100 μL acetic acid and 50 μL of 10% TFA. The mixture was analyzed by HPLC, and the obtained profile was compared to that of a control not subjected to protease treatment. Proteolytic degradation was considered insignificant if the change in macrocycle peak area was less than 10% (within the precision limit of the method). The HPLC separation was performed on a Jupiter C18 300A column (5 μm, 250 × 4.6 mm; Phenomenex) equilibrated in 10% (v/v) acetonitrile in 0.1% (v/v) TFA and eluted with a linear gradient of 99% (v/v) acetonitrile in 0.1% (v/v) TFA.

Cell-Based Assays. Compound cytotoxicity was evaluated on nontumor human dermal fibroblasts and peripheral blood mononuclear cells. Cells were seeded in 384-well plates at a density of 2,000−50,000 cells per well, and the tested compounds were added at varying concentrations. The cells were incubated at 37 °C, 5% CO_2 for 72 h, after which CellTiter-Glo (Promega) detection reagent was added. The luminescence signal of the compound-treated cells was compared to that of an untreated control set to 100% cell viability. The transport potential of the tested compounds across cell membranes was evaluated with a Caco-2 permeability assay (BD BioCoat HTS Caco-2 assay system, BD Biosciences) as previously described.[47] Inhibitors were applied at 5 μM concentration to the apical side of the Caco-2 cell monolayer. After 3 h, the concentration of the compounds was determined in both the apical and basolateral compartments using liquid chromatography−mass spectrometry, and apparent permeability (P_{app}) was calculated.

Crystal Structure Determination. Preparation and crystallization of complexes of human CatD with inhibitors were as follows. Purified human CatD (7.5 mg mL^{-1}) was incubated for 30 min at room temperature with a twofold molar excess of macrocyclic inhibitor (applied as DMSO solution) in 25 mM Tris-HCl, pH 7.0 (the final DMSO concentration did not exceed 7%). The mixture was centrifuged for 5 min at 16,000g. Crystals were grown at 20 °C by the hanging-drop vapor-diffusion method. Drops were equilibrated over a precipitant solution of 0.1 M Bis-Tris, pH 5.5, containing 25% PEG 3350, and 0.2 M Li_2SO_4. The ratio of protein to precipitant solution was 2:1 for CatD-**9** and CatD-**33** complexes and 1:1 for CatD-**14**. The crystals were flash-cooled in liquid nitrogen with the precipitant solution as self-cryoprotectant. Diffraction data were collected at 100 K on MX 14.1 operated by the Joint Berlin MX-Laboratory at the BESSY II electron storage ring in Berlin-Adlershof, Germany. All diffraction data were processed using the XDS suite of programs. The structures of the complexes were solved by molecular replacement with the program MOLREP[48] using uncomplexed CatD (PDB code 1LYA) as a search model. Model refinement was carried out using the REFMAC 5.2 program[49] from the CCP4 package,[50] with manual adjustments in Coot.[51] The quality of the final model was validated with MolProbity.[52] The final refinement statistics are summarized in Supporting Information Table 4. All figures showing structural representations were prepared with the program PyMOL (Schrodinger LLC). Atomic contacts between inhibitors and protein moiety were generated using CCP4 Contact. Crystal parameters and data collection statistics are given in Supporting Information Table 4.

Interaction Energy Calculations. The crystal structures were examined by the hybrid QM/SQM method. Hydrogen atoms were added at the authentic protonation state, and water molecule W14

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

was included in the calculations. Noncrystallographic complexes were built in multiple orientations based on CatD-**14** and CatD-**9** crystal structures.

Structure Preparation. Hydrogen atoms were added by the Reduce and Leap programs in AMBER 14.[53] PROPKA 3.0[54] was used to assign protonation states. Aspartates, glutamates, lysines, arginines, and histidines were charged, except for Glu5, Asp50, Glu180, Asp231, and Asp323. Hydrogen atoms were relaxed by annealing from 600 to 0 K at the MM level in AMBER 14. The FF14SB force field was used for the protein, whereas the GAFF force field was used for the ligand. The cooling runs were 4 ps long with 1 fs step and Berendsen thermostat used.

QM/SQM Setup. The QM part comprised residues located within 2.5 Å of inhibitor **9** in the CatD-**9** crystal structure. The QM part was treated at the DFT-D3/BLYP/TZVPP level for energy calculations and at the DFT-D3/BLYP/DZVP level for optimization.[55] The rest of the system was treated at the PM6-D3H4X level.[56,57] The environment was described by the COSMO implicit solvent model.[58] The coupling between QM and SQM was performed with Cuby4,[59] which calls Turbomole 7.0[60] and MOPAC[61] for QM and SQM, respectively. Residues farther than 6 Å from inhibitor **9** in the CatD-**9** crystal structure were frozen during the optimization.

Scoring Function. The score was computed as a sum of gas-phase interaction energy (ΔE_{int}), the interaction solvation free energy ($\Delta\Delta G_{solv}$), and the change in conformational "free" energy of the ligand ($\Delta G'^{W}_{conf}(L)$).[62,63] ΔE_{int} was computed using the QM/SQM methodology described above. Two methods were used for solvation free energy calculations: PM6/COSMO implemented in MOPAC for the protein and SMD/HF/6-31G*[64] implemented in Gaussian09[65] for the ligand. $\Delta G'^{W}_{conf}(L)$ was computed as the "free" energy change between the ligand structure in the optimized protein−ligand complex and the ligand structure optimized in solution by the same method. The "free" energy was defined as the sum of the gas-phase DFT-D3/TPSS/TZVPP energy with the SMD/HF/6-31G* solvation free energy for the $\Delta G'^{W}_{conf}(L)$ term.

Glycine Scan and Inhibitor Fragmentation. The energy contributions of individual amino acid residues were calculated as the difference between the interaction "free" energy ($\Delta G'$) with the wild-type amino acid and the $\Delta G'$ with the mutated glycine residue. For the purposes of inhibitor fragmentation, selected C−C bonds of the inhibitors were cut and capped with hydrogens, and $\Delta G'$ between the ligand fragment and the whole protein was computed. The $\Delta G'$ values were determined using the QM/SQM methodology described above (DFT-D3/TPSS/TZVPP for QM and PM6-D3H4X for SQM combined with the COSMO implicit solvent model).

Noncrystallographic Complexes. The complex of CatD with **22** was built based on the CatD-**14** crystal structure. Complexes with inhibitors **8**, **24**, **25**, and **31** were built based on the CatD-**9** crystal structure. The ligand modifications were built manually using the PyMOL program in multiple orientations. All complexes were optimized using the hybrid QM/SQM methodology, and the most stable complexes were used for scoring (as described above).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jmedchem.9b01351.

Physicochemical parameters of pepstatin A (**1**) and the most potent macrocyclic inhibitors with subnanomolar and single-digit nanomolar inhibition of CatD; cytotoxicity analysis of model macrocyclic inhibitors and pepstatin A (**1**); correlation of computed scores plotted against experimental binding free energy for selected macrocyclic inhibitors; inhibition kinetics for the macrocyclic inhibitor 9 and human CatD; list of atomic contacts between macrocyclic inhibitors and CatD; X-ray data collection and refinement statistics; and analytical RP-HPLC traces for the key compounds (PDF)

Molecular formula strings (CSV)

Protien data bank for CatD-**9** complex (PDB)

Protien data bank for CatD-**14** complex (PDB)

Protien data bank for CatD-**33** complex (PDB)

Accession Codes

The structures presented in this article have been deposited in the Protein Data Bank under accession codes 6QCB, 6QBG, and 6QBH for the CatD-**9**, CatD-**14**, and CatD-**33** complexes, respectively. Authors will release the atomic coordinates and experimental data upon article publication.

■ AUTHOR INFORMATION

Corresponding Authors

Pavel Majer − *Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Praha 6, Czech Republic*; Email: majer@uochb.cas.cz

Michael Mareš − *Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Praha 6, Czech Republic*; orcid.org/0000-0002-0847-5022; Email: mares@uochb.cas.cz

Authors

Radka Houštecká − *Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Praha 6, Czech Republic; First Faculty of Medicine, Charles University, 12108 Praha 2, Czech Republic*

Martin Hadzima − *Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Praha 6, Czech Republic; Department of Organic Chemistry, Faculty of Science, Charles University, 12800 Praha 2, Czech Republic*

Jindřich Fanfrlík − *Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Praha 6, Czech Republic*

Jiří Brynda − *Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Praha 6, Czech Republic*

Lenka Pallová − *Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Praha 6, Czech Republic*

Iva Hánová − *Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Praha 6, Czech Republic; Department of Biochemistry, Faculty of Science, Charles University, 12800 Praha 2, Czech Republic*; orcid.org/0000-0002-7173-8200

Helena Mertlíková-Kaiserová − *Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Praha 6, Czech Republic*

Martin Lepšík − *Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Praha 6, Czech Republic*

Martin Horn − *Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Praha 6, Czech Republic*; orcid.org/0000-0001-9110-2018

Martin Smrčina − *Tucson Research Center, Icagen Inc., Oro Valley, Arizona 85755, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jmedchem.9b01351

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by project ChemBioDrug CZ.02.1.01/0.0/0.0/16_019/0000729 from the European Regional Development Fund (OP RDE), grant 15-18929S from the Czech Science Foundation, project InterBioMed LO1302 from the Ministry of Education of the Czech Republic, grant 802216 from the Charles University Grant Agency, Gilead Sciences and IOCB Research Center, and institutional project RVO 61388963. We would like to thank Radko Souček for amino acid analysis and solubility analysis, and Pavlína Řezáčová and Taťána Majerová (Institute of Organic Chemistry and Biochemistry, C.A.S., Prague, Czech Republic) for materials for the HIV-1 protease assay. We are also very grateful to Jaroslav Feyereisl and voluntary donors from the Institute for the Care of Mother and Child (Prague, Czech Republic) for placental tissue.

■ ABBREVIATIONS

Abz, aminobenzoic acid; AMC, 7-amino-4-methylcoumarin; CatD, cathepsin D; Iva, isovaleryl; Sta, statine; TSTU, *N*,*N*,*N*′,*N*′-tetramethyl-*O*-(*N*-succinimidyl)uronium tetrafluoroborate; Z, *N*-carbobenzyloxy

■ REFERENCES

(1) Fusek, M.; Mareš, M.; Větvička, V. Cathepsin D. In *Handbook of Proteolytic Enzymes*; Rawlings, N., Salvesen, G., Eds.; Academic Press: Oxford, 2013; pp 54−63.

(2) Vidoni, C.; Follo, C.; Savino, M.; Melone, M. A. B.; Isidoro, C. The Role of Cathepsin D in the Pathogenesis of Human Neurodegenerative Disorders. *Med. Res. Rev.* **2016**, *36*, 845−870.

(3) Aghdassi, A. A.; John, D. S.; Sendler, M.; Weiss, F. U.; Reinheckel, T.; Mayerle, J.; Lerch, M. M. Cathepsin D Regulates Cathepsin B Activation and Disease Severity Predominantly in Inflammatory Cells during Experimental Pancreatitis. *J. Biol. Chem.* **2018**, *293*, 1018−1029.

(4) Liaudet-Coopman, E.; Beaujouin, M.; Derocq, D.; Garcia, M.; Glondu-Lassis, M.; Laurent-Matha, V.; Prébois, C.; Rochefort, H.; Vignon, F. Cathepsin D: Newly Discovered Functions of a Long-Standing Aspartic Protease in Cancer and Apoptosis. *Cancer Lett.* **2006**, *237*, 167−179.

(5) Benes, P.; Vetvicka, V.; Fusek, M. Cathepsin D—Many Functions of One Aspartic Protease. *Crit. Rev. Oncol. Hematol.* **2008**, *68*, 12−28.

(6) Foekens, J. A.; Look, M. P.; Bolt-de Vries, J. B.; Meijer-van Gelder, M. E.; van Putten, W. L. J.; Klijn, J. G. M. Cathepsin-D in Primary Breast Cancer: Prognostic Evaluation Involving 2810 Patients. *Br. J. Cancer* **1999**, *79*, 300−307.

(7) Ferrandina, G.; Scambia, G.; Bardelli, F.; Benedetti Panici, P.; Mancuso, S.; Messori, A. Relationship between Cathepsin-D Content and Disease-Free Survival in Node-Negative Breast Cancer Patients: A Meta-Analysis. *Br. J. Cancer* **1997**, *76*, 661−666.

(8) Ashraf, Y.; Mansouri, H.; Laurent-Matha, V.; Alcaraz, L. B.; Roger, P.; Guiu, S.; Derocq, D.; Robin, G.; Michaud, H.-A.; Delpech, H.; Jarlier, M.; Pugnière, M.; Robert, B.; Puel, A.; Martin, L.; Landomiel, F.; Bourquard, T.; Achour, O.; Fruitier-Arnaudin, I.; Pichard, A.; Deshayes, E.; Turtoi, A.; Poupon, A.; Simony-Lafontaine, J.; Boissière-Michot, F.; Pirot, N.; Bernex, F.; Jacot, W.; du Manoir, S.; Theillet, C.; Pouget, J.-P.; Navarro-Teulon, I.; Bonnefoy, N.; Pèlegrin, A.; Chardès, T.; Martineau, P.; Liaudet-Coopman, E. Immunotherapy of Triple-Negative Breast Cancer with Cathepsin D-Targeting Antibodies. *J. Immunother. Cancer* **2019**, *7*, 29.

(9) Masson, O.; Bach, A.-S.; Derocq, D.; Prébois, C.; Laurent-Matha, V.; Pattingre, S.; Liaudet-Coopman, E. Pathophysiological Functions of Cathepsin D: Targeting Its Catalytic Activity versus Its Protein Binding Activity? *Biochimie* **2010**, *92*, 1635−1643.

(10) Wood, J. M.; Maibaum, J.; Rahuel, J.; Grütter, M. G.; Cohen, N.-C.; Rasetti, V.; Rüger, H.; Göschke, R.; Stutz, S.; Fuhrer, W.; Schilling, W.; Rigollier, P.; Yamaguchi, Y.; Cumin, F.; Baum, H.-P.; Schnell, C. R.; Herold, P.; Mah, R.; Jensen, C.; O'Brien, E.; Stanton, A.; Bedigian, M. P. Structure-Based Design of Aliskiren, a Novel Orally Effective Renin Inhibitor. *Biochem. Biophys. Res. Commun.* **2003**, *308*, 698−705.

(11) Ghosh, A. K.; Osswald, H. L.; Prato, G. Recent Progress in the Development of HIV-1 Protease Inhibitors for the Treatment of HIV/AIDS. *J. Med. Chem.* **2016**, *59*, 5172−5208.

(12) Ersmark, K.; Samuelsson, B.; Hallberg, A. Plasmepsins as Potential Targets for New Antimalarial Therapy. *Med. Res. Rev.* **2006**, *26*, 626−666.

(13) Coimbra, J. R. M.; Marques, D. F. F.; Baptista, S. J.; Pereira, C. M. F.; Moreira, P. I.; Dinis, T. C. P.; Santos, A. E.; Salvador, J. A. R. Highlights in BACE1 Inhibitors for Alzheimer's Disease Treatment. *Front. Chem.* **2018**, *6*, 178.

(14) Kick, E. K.; Roe, D. C.; Geoffrey Skillman, A.; Liu, G.; Ewing, T. J. A.; Sun, Y.; Kuntz, I. D.; Ellman, J. A. Structure-Based Design and Combinatorial Chemistry Yield Low Nanomolar Inhibitors of Cathepsin D. *Chem. Biol.* **1997**, *4*, 297−307.

(15) Srivastava, V.; Saxena, H. O.; Shanker, K.; Kumar, J. K.; Luqman, S.; Gupta, M. M.; Khanuja, S. P. S.; Negi, A. S. Synthesis of Gallic Acid Based Naphthophenone Fatty Acid Amides as Cathepsin D Inhibitors. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4603−4608.

(16) Grädler, U.; Czodrowski, P.; Tsaklakidis, C.; Klein, M.; Werkmann, D.; Lindemann, S.; Maskos, K.; Leuthner, B. Structure-Based Optimization of Non-Peptidic Cathepsin D Inhibitors. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 4141−4150.

(17) Lee, C. E.; Kick, E. K.; Ellman, J. A. General Solid-Phase Synthesis Approach to Prepare Mechanism-Based Aspartyl Protease Inhibitor Libraries. Identification of Potent Cathepsin D Inhibitors. *J. Am. Chem. Soc.* **1998**, *120*, 9735−9747.

(18) Knight, C. G.; Barrett, A. J. Interaction of Human Cathepsin D with the Inhibitor Pepstatin. *Biochem. J.* **1976**, *155*, 117−125.

(19) Majer, P.; Collins, J. R.; Gulnik, S. V.; Erickson, J. W. Structure-Based Subsite Specificity Mapping of Human Cathepsin D Using Statine-Based Inhibitors. *Protein Sci.* **1997**, *6*, 1458−1466.

(20) Nussbaumerová, M.; Srp, J.; Máša, M.; Hradilek, M.; Šanda, M.; Reiniš, M.; Horn, M.; Mareš, M. Single- and Double-Headed Chemical Probes for Detection of Active Cathepsin D in a Cancer Cell Proteome. *ChemBioChem* **2010**, *11*, 1538−1541.

(21) Horn, M.; Nussbaumerová, M.; Šanda, M.; Kovářová, Z.; Srba, J.; Franta, Z.; Sojka, D.; Bogyo, M.; Caffrey, C. R.; Kopáček, P.; Mareš, M. Hemoglobin Digestion in Blood-Feeding Ticks: Mapping a Multipeptidase Pathway by Functional Proteomics. *Chem. Biol.* **2009**, *16*, 1053−1063.

(22) Tumminello, F. M.; Bernacki, R. J.; Gebbia, N.; Leto, G. Pepstatins: Aspartic Proteinase Inhibitors Having Potential Therapeutic Applications. *Med. Res. Rev.* **1993**, *13*, 199−208.

(23) Agarwal, N. S.; Rich, D. H. Inhibition of Cathepsin D by Substrate Analogues Containing Statine and by Analogues of Pepstatin. *J. Med. Chem.* **1986**, *29*, 2519−2524.

(24) Xie, D.; Gulnik, S.; Collins, L.; Gustchina, E.; Suvorov, L.; Erickson, J. W. Dissection of the pH Dependence of Inhibitor Binding Energetics for an Aspartic Protease: Direct Measurement of the Protonation States of the Catalytic Aspartic Acid Residues. *Biochemistry* **1997**, *36*, 16166−16172.

(25) Cummings, M. D.; Sekharan, S. Structure-Based Macrocycle Design in Small-Molecule Drug Discovery and Simple Metrics To Identify Opportunities for Macrocyclization of Small-Molecule Ligands. *J. Med. Chem.* **2019**, *62*, 6843−6853.

(26) Jouin, P.; Castro, B.; Nisato, D. Stereospecific Synthesis of N-Protected Statine and Its Analogues via Chiral Tetramic Acid. *J. Chem. Soc., Perkin Trans. 1* **1987**, 1177−1182.

(27) Smrcina, M.; Majer, P.; Majerová, E.; Guerassina, T. A.; Eissenstat, M. A. Facile Stereoselective Synthesis of γ-Substituted γ-Amino Acids from the Corresponding α-Amino Acids. *Tetrahedron* **1997**, *53*, 12867−12874.

https://dx.doi.org/10.1021/acs.jmedchem.9b01351
J. Med. Chem. 2020, 63, 1576−1596

(28) Nishikawa, K.; Yoshimi, Y.; Maeda, K.; Morita, T.; Takahashi, I.; Itou, T.; Inagaki, S.; Hatanaka, M. Radical Photocyclization Route for Macrocyclic Lactone Ring Expansion and Conversion to Macrocyclic Lactams and Ketones. *J. Org. Chem.* **2013**, *78*, 582−589.
(29) Rich, D. H.; Sun, E. T. O.; Ulm, E. Synthesis of Analogues of the Carboxyl Protease Inhibitor Pepstatin. Effects of Structure on Inhibition of Pepsin and Renin. *J. Med. Chem.* **1980**, *23*, 27−33.
(30) Baldwin, E. T.; Bhat, T. N.; Gulnik, S.; Hosur, M. V.; Sowder, R. C.; Cachau, R. E.; Collins, J.; Silva, A. M.; Erickson, J. W. Crystal Structures of Native and Inhibited Forms of Human Cathepsin D: Implications for Lysosomal Targeting and Drug Design. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6796−6800.
(31) Giordanetto, F.; Kihlberg, J. Macrocyclic Drugs and Clinical Candidates: What Can Medicinal Chemists Learn from Their Properties? *J. Med. Chem.* **2014**, *57*, 278−295.
(32) Marsault, E.; Peterson, M. L. Macrocycles Are Great Cycles: Applications, Opportunities, and Challenges of Synthetic Macrocycles in Drug Discovery. *J. Med. Chem.* **2011**, *54*, 1961−2004.
(33) Mallinson, J.; Collins, I. Macrocycles in New Drug Discovery. *Future Med. Chem.* **2012**, *4*, 1409−1438.
(34) Tyndall, J.; Fairlie, D. Macrocycles Mimic The Extended Peptide Conformation Recognized By Aspartic, Serine, Cysteine and Metallo Proteases. *Curr. Med. Chem.* **2001**, *8*, 893−907.
(35) Madala, P. K.; Tyndall, J. D. A.; Nall, T.; Fairlie, D. P. Update 1 of: Proteases Universally Recognize Beta Strands in Their Active Sites. *Chem. Rev.* **2010**, *110*, PR1−PR31.
(36) Fusek, M.; Baudyš, M.; Metcalf, P. Purification and Crystallization of Human Cathepsin D. *J. Mol. Biol.* **1992**, *226*, 555−557.
(37) Hánová, I.; Brynda, J.; Houštecká, R.; Alam, N.; Sojka, D.; Kopáček, P.; Marešová, L.; Vondrášek, J.; Horn, M.; Schueler-Furman, O.; Mareš, M. Novel Structural Mechanism of Allosteric Regulation of Aspartic Peptidases via an Evolutionarily Conserved Exosite. *Cell Chem. Biol.* **2018**, *25*, 318−329.e4.
(38) Máša, M.; Marešová, L.; Vondrášek, J.; Horn, M.; Ježek, J.; Mareš, M. Cathepsin D Propeptide: Mechanism and Regulation of Its Interaction with the Catalytic Core. *Biochemistry* **2006**, *45*, 15474−15482.
(39) Baell, J. B.; Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010**, *53*, 2719−2740.
(40) Žebrakovská, I.; Máša, M.; Srp, J.; Horn, M.; Vávrová, K.; Mareš, M. Complex Modulation of Peptidolytic Activity of Cathepsin D by Sphingolipids. *Biochim. Biophys. Acta, Mol. Cell Biol. Lipids* **2011**, *1811*, 1097−1104.
(41) Uhlíková, T.; Konvalinka, J.; Pichová, I.; Souček, M.; Kräusslich, H.-G.; Vondrášek, J. A Modular Approach to HIV-1 Proteinase Inhibitor Design. *Biochem. Biophys. Res. Commun.* **1996**, *222*, 38−43.
(42) Andrau, D.; Dumanchin-Njock, C.; Ayral, E.; Vizzavona, J.; Farzan, M.; Boisbrun, M.; Fulcrand, P.; Hernandez, J.-F.; Martinez, J.; Lefranc-Jullien, S.; Checler, F. BACE1- and BACE2-Expressing Human Cells. *J. Biol. Chem.* **2003**, *278*, 25859−25866.
(43) Holzman, T. F.; Chung, C. C.; Edalji, R.; Egan, D. A.; Gubbins, E. J.; Rueter, A.; Howard, G.; Yang, L. K.; Pederson, T. M.; Krafft, G. A.; Wang, G. T. Recombinant Human Prorenin from CHO Cells: Expression and Purification. *J. Protein Chem.* **1990**, *9*, 663−672.
(44) Salát, J.; Paesen, G. C.; Řezáčová, P.; Kotsyfakis, M.; Kovářová, Z.; Šanda, M.; Majtán, J.; Grunclová, L.; Horká, H.; Andersen, J. F.; Brynda, J.; Horn, M.; Nunn, M. A.; Kopáček, P.; Kopecký, J.; Mareš, M. Crystal Structure and Functional Characterization of an Immunomodulatory Salivary Cystatin from the Soft Tick Ornithodoros moubata. *Biochem. J.* **2010**, *429*, 103−112.
(45) Srp, J.; Nussbaumerová, M.; Horn, M.; Mareš, M. Digestive Proteolysis in the Colorado Potato Beetle, Leptinotarsa decemlineata: Activity-Based Profiling and Imaging of a Multipeptidase Network. *Insect Biochem. Mol. Biol.* **2016**, *78*, 1−11.
(46) Jílková, A.; Horn, M.; Řezáčová, P.; Marešová, L.; Fajtová, P.; Brynda, J.; Vondrášek, J.; McKerrow, J. H.; Caffrey, C. R.; Mareš, M. Activation Route of the Schistosoma mansoni Cathepsin B1 Drug Target: Structural Map with a Glycosaminoglycan Switch. *Structure* **2014**, *22*, 1786−1798.
(47) Šmídková, M.; Dvořáková, A.; Tloust'ová, E.; Česnek, M.; Janeba, Z.; Mertlíková-Kaiserová, H. Amidate Prodrugs of 9-[2-(Phosphonomethoxy)Ethyl]Adenine as Inhibitors of Adenylate Cyclase Toxin from Bordetella pertussis. *Antimicrob. Agents Chemother.* **2014**, *58*, 664−671.
(48) Vagin, A.; Teplyakov, A. An Approach to Multi-Copy Search in Molecular Replacement. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2000**, *56*, 1622−1624.
(49) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53*, 240−255.
(50) Bailey, S. The Ccp4 Suite - Programs for Protein Crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1994**, *50*, 760−763.
(51) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126−2132.
(52) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 12−21.
(53) Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Kollman, P. A. *Amber 14*; University of California: San Francisco, 2014.
(54) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525−537.
(55) Hostaš, J.; Řezáč, J. Accurate DFT-D3 Calculations in a Small Basis Set. *J. Chem. Theory Comput.* **2017**, *13*, 3575−3585.
(56) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements. *J. Mol. Model.* **2007**, *13*, 1173−1213.
(57) Řezáč, J.; Hobza, P. Advanced Corrections of Hydrogen Bonding and Dispersion for Semiempirical Quantum Mechanical Methods. *J. Chem. Theory Comput.* **2012**, *8*, 141−151.
(58) Klamt, A.; Schüürmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799−805.
(59) Řezáč, J. Cuby: An Integrative Framework for Computational Chemistry. *J. Comput. Chem.* **2016**, *37*, 1230−1237.
(60) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic Structure Calculations on Workstation Computers: The Program System Turbomole. *Chem. Phys. Lett.* **1989**, *162*, 165−169.
(61) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods IV: Extension of MNDO, AM1, and PM3 to More Main Group Elements. *J. Mol. Model.* **2004**, *10*, 155−164.
(62) Lepšík, M.; Řezáč, J.; Kolář, M.; Pecina, A.; Hobza, P.; Fanfrlík, J. The Semiempirical Quantum Mechanical Scoring Function for in Silico Drug Design. *Chempluschem* **2013**, *78*, 921−931.
(63) Fanfrlík, J.; Bronowska, A. K.; Řezáč, J.; Přenosil, O.; Konvalinka, J.; Hobza, P. A Reliable Docking/Scoring Scheme Based on the Semiempirical Quantum Mechanical PM6-DH2 Method Accurately Covering Dispersion and H-Bonding: HIV-1 Protease with 22 Ligands. *J. Phys. Chem. B* **2010**, *114*, 12666−12678.
(64) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a

SUPPORTING INFORMATION

Biomimetic macrocyclic inhibitors of human cathepsin D: structure-activity relationship and binding mode analysis

Radka Houštecká, Martin Hadzima, Jindřich Fanfrlík, Jiří Brynda, Lenka Pallová, Iva Hánová, Helena Mertlíková-Kaiserová, Martin Lepšík, Martin Horn, Martin Smrčina, Pavel Majer, Michael Mareš

Table of contents:

Supplementary Table 1. Physicochemical parameters of pepstatin A (1) and the most potent macrocyclic inhibitors with subnanomolar and single-digit nanomolar inhibition of CatD ……..…S2
Supplementary Table 2. Cytotoxicity analysis of model macrocyclic inhibitors and pepstatin A (**1**)……...S2
Supplementary Figure 1. Correlation of computed scores plotted against experimental binding free energy for selected macrocyclic inhibitors……..……..……..S3
Supplementary Figure 2. Inhibition kinetics for the macrocyclic inhibitor **9** and human CatD…...S3
Supplementary Table 3. List of atomic contacts between macrocyclic inhibitors and CatD……...S4
Supplementary Table 4. X-ray data collection and refinement statistics……..S5
Supplementary Figure 3. Analytical RP-HPLC traces for the key compounds…….....S6

Supplementary Table 1. Physicochemical parameters of pepstatin A (1) and the most potent macrocyclic inhibitors with subnanomolar and single-digit nanomolar inhibition of CatD

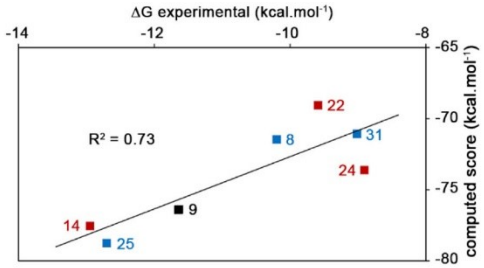
Compound	MW	nON[a]	nOHNH[b]	logP	tPSA (Å²)[c]	NRotB[d]
1	685.9	14	8	1.43	233.26	22
9	487.7	7	4	3.61	107.53	3
11	501.7	7	4	3.99	107.50	3
14	537.7	7	4	4.85	107.53	3
15	523.7	7	4	3.59	107.53	3
16	556.6	7	4	4.85	107.53	3
17	505.7	7	4	3.74	107.53	3
18	494.0	7	4	4.20	107.53	3
25	536.7	8	4	3.34	119.89	4
26	501.7	7	4	4.00	107.53	4
27	527.7	7	4	4.24	107.53	3
33	503.7	8	4	3.81	116.76	3

[a] Number of hydrogen bond acceptors. [b] Number of hydrogen bond donors. [c] Topological polar surface area. [d] Number of rotatable bonds. Physicochemical parameters were determined from https://www.molinspiration.com/cgi-bin/properties and ALOGPS2.1 software from http://www.vcclab.org/web/alogps/.

Supplementary Table 2. Cytotoxicity analysis of representative macrocyclic inhibitors and pepstatin A (1)

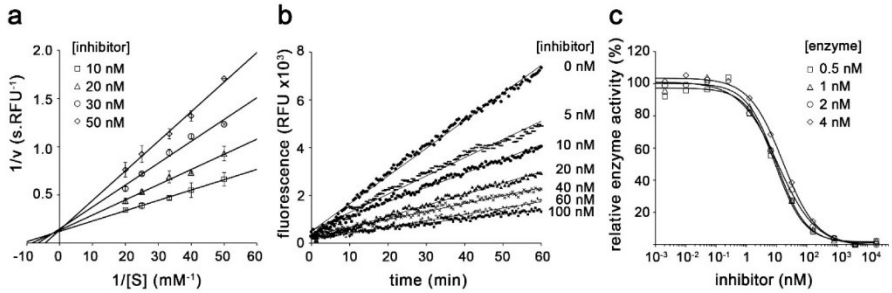
Compound	Concentration [µM]	Cell viability (%)[a]	
		NHDF[b]	PBMC[c]
1	0	100	100
	1	92	110
	10	98	99
	25	97	94
	50	96	94
	100	88	84
25	0	100	100
	1	97	110
	10	100	99
	25	94	94
	50	91	94
	100	82	84
33	0	100	100
	1	97	92
	10	102	107
	25	100	104
	50	98	92
	100	95	85

[a] Results of *in vitro* cytotoxicity assays on non-cancerous cell lines are expressed as percent viability compared to the untreated control. Standard deviation values are within 10% of the mean of tetraplicates. [b] Normal human dermal fibroblasts. [c] Human peripheral blood mononuclear cells.



Compound	Score	ΔE_{int}	$\Delta\Delta G_{solv}$	$\Delta G'^{W}_{conf}(L)$
14	-77.5	-165.9	82.1	6.2
25	-78.8	-175.2	93.1	3.3
9	-76.4	-162.6	82.8	3.4
8	-71.5	-157.9	82.8	3.6
22	-69.1	-158.2	83.0	6.1
31	-71.1	-166.3	87.7	7.6
24	-73.6	-158.9	81.4	4.0

Supplementary Figure 1. Correlation of computed scores plotted against experimental binding free energy for selected macrocyclic inhibitors. Compounds highlighted in red and blue belong to the groups listed in Tables 3a and 3b, respectively; **9** (black) belongs to both groups. All energy values in the table are expressed in kcal.mol^{-1}; the score is a sum of the gas-phase interaction energy (ΔE_{int}), the interaction solvation free energy ($\Delta\Delta G_{solv}$), and the change in conformational 'free' energy of the ligand ($\Delta G'^{W}_{conf}(L)$).



Supplementary Figure 2. Inhibition kinetics for the macrocyclic inhibitor **9** and human CatD. (**a**) The double reciprocal plot with various inhibitor and substrate concentrations reveals a competitive mode (the lines intersect on the 1/v axis). (**b**) Fast-binding inhibitor kinetics were identified by continuous activity monitoring for 1 h at different inhibitor concentrations. (**c**) IC_{50} values are independent of enzyme concentration, indicating that the inhibitor is not 'tight-binding.' CatD activity was measured in a kinetic assay with a FRET-peptide substrate (S). Error bars depict the mean ± SE from triplicate measurements.

Supplementary Table 3. List of atomic contacts between macrocyclic inhibitors and CatD. Analysis was performed using the program CCP4. The distance cutoffs were set to 3.3 Å and 4.0 Å for hydrogen bonds (H-bonds) and van der Waals interactions, respectively. For each CatD residue, the total number of contacts, including H-bonds, is indicated. (**a**) CatD residues interacting at individual positions (P1, P2' and the bridge) of three macrocyclic inhibitors. (**b**) CatD residues interacting with individual pepstatin A positions (P4 to P3') analyzed for PDB structure 1LYB.

a

		compound 9		compound 14		compound 33	
	residue	contacts	H-bonds	contacts	H-bonds	contacts	H-bonds
P1	Asp 33	10	1	10	1	10	1
	Gly 35	5		5		7	
	Tyr 78	12		11		11	
	Gly 79	1	1	1	1	1	1
	Ser 80	3		6		5	
	Thr 125	0		1		0	
	Phe 126	2		1		2	
	Phe 131	2		1		2	
	Ile 134	5		2		4	
	Asp 231	5	1	5	1	5	1
	Gly 233	9	1	9	1	7	1
	Thr 234	1		1		0	
P2'	Gly 35	5	1	4	1	4	1
	Ser 36	3		2		2	
	Ile 76	1		1		0	
	His 77	3		3		3	
	Tyr 78	2		3		2	
	Ile 142	1		1		1	
	Tyr 205	4	1	4	1	4	1
bridge	His 77	3	1	4	1	4	1
	Tyr 78	1		1		3	
	Gly 79	6	1	4	1	5	1
	Ser 80	5	2	7	2	6	1
	Gly 233	2		0		0	
	Thr 234	3		3		3	
	Met 307	0		0		1	
	Met 309	3		2		3	
	Ile 320	2		1		2	
sum		99	10	93	10	97	9

b

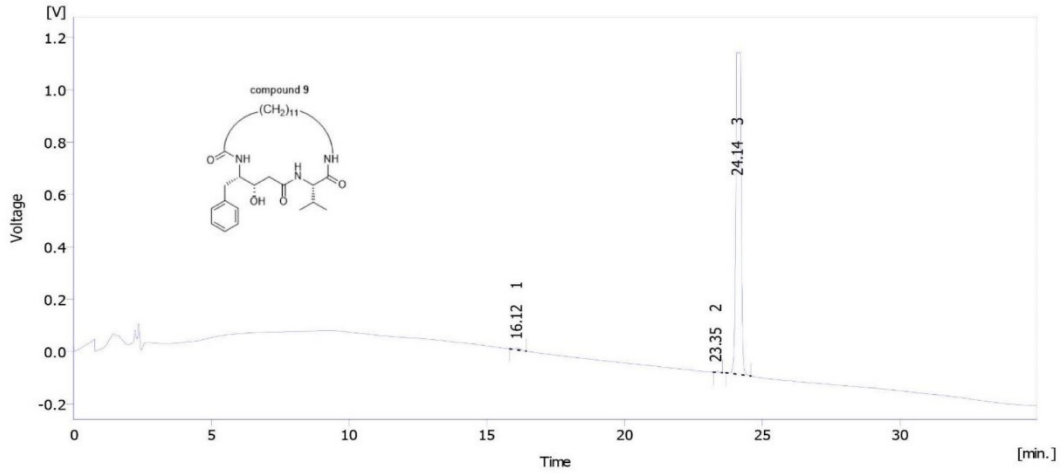
		pepstatin	
	residue	contacts	H-bonds
P4 (Iva)	Ser 235	3	1
	Glu 260	1	
P3 (Val)	Ser 80	3	
	Gly 233	3	
	Thr 234	2	
	Ser 235	5	1
P2 (Val)	Gly 79	4	
	Ser 80	10	3
	Gly 233	3	
	Thr 234	6	
	Ile 320	1	
P1 (Sta)	Val 31	1	
	Asp 33	8	1
	Gly 35	3	
	Tyr 78	6	
	Gly 79	2	1
	Ser 80	1	
	Ile 134	1	
	Asp 231	5	1
	Gly 233	8	1
	Thr 234	2	
P2' (Ala)	Gly 35	4	1
	Ser 36	1	
	His 77	1	
	Tyr 78	2	
	Tyr 205	4	1
P3' (Sta)	His 77	5	
	Tyr 205	4	
	Ile 311	2	
	Ile 320	1	
sum		102	11

Supplementary Table 4. X-ray data collection and refinement statistics.

	CatD-9 6QCB	CatD-14 6QBG	CatD-33 6QBH
Data collection			
Space group	$C222_1$	$C222_1$	$C222_1$
Cell dimensions			
a, *b*, *c* (Å)	80.35 94.26 117.68	80.37 94.14 117.51	80.24 94.14 117.58
α, β, γ (°)	90.00 90.00 90.00	90.00 90.00 90.00	90.00 90.00 90.00
Resolution (Å)	50.0–1.55 (1.60–1.55)	50.0–1.8 (1.85–1.80)	50.0–1.85 (1.90–1.85)
*R*merge (%)	7.8 (95.3)	11.8 (132.4)	10.9 (96.4)
$I / \sigma I$	12.5 (1.02)	7.9 (0.97)	7.9 (1.10)
Completeness (%)	97.8 (88.5)	98.8 (98.3)	98.6 (97.4)
Multiplicity	3.5 (2.8)	3.6 (3.6)	3.6 (3.4)
$CC_{1/2}$	99.7 (62.2)	99.6 (56.4)	99.6 (51.2)
Refinement			
Resolution (Å)	43.75–1.55 (1.60–1.55)	61.1–1.8 (1.85–1.79)	47.07–1.85 (1.90–1.85)
No. reflections	62,404	39,934	32,236
*R*work	0.166 (0.274)	0.194 (0.434)	0.177 (0.311)
*R*free	0.189 (0.275)	0.247 (0.416)	0.206 (0.293)
No. atoms			
Protein	2681	2617	2628
Inhibitor	35	78	36
Water	323	252	250
No. molecules in AU	1	1	1
B-factors			
Protein	21.1	28.2	19.3
Inhibitor	19.4	32.4	17.3
Water	31.4	33.6	26.4
R.m.s deviations			
Bond lengths (Å)	0.014	0.015	0.011
Bond angles (°)	1.740	1.833	1.619
MolProbity analysis			
Ramachandran outlier (%)	0	0	0
Ramachandran favoured (%)	99.1	98.2	98.2
MolProbity score	1.18	0.91	1.07

Supplementary Figure 3. Analytical RP-HPLC traces for the key compounds. Experimental details are described in the General procedures in the Experimental section.

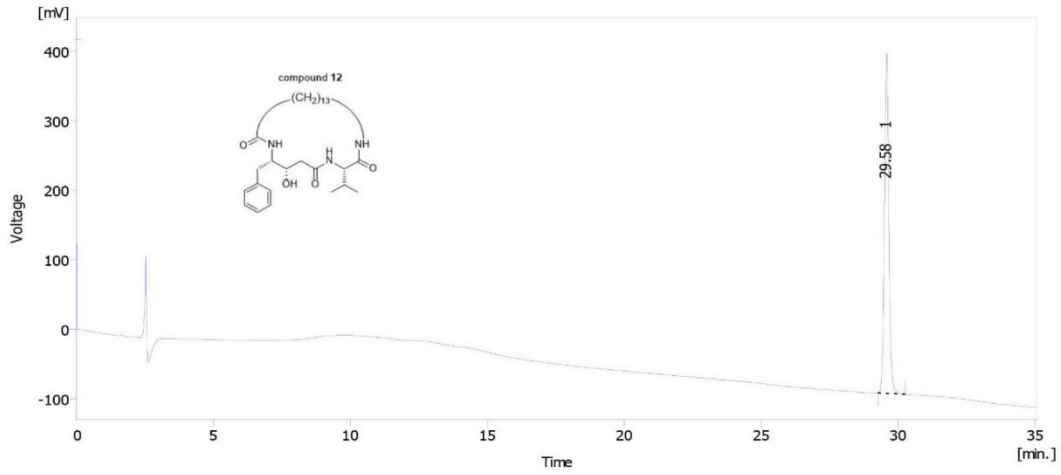
Compound 9



Result Table

	Reten. Time [min]	Area [%]
1	16.120	0.6
2	23.352	0.2
3	24.140	99.2
	Total	100.0

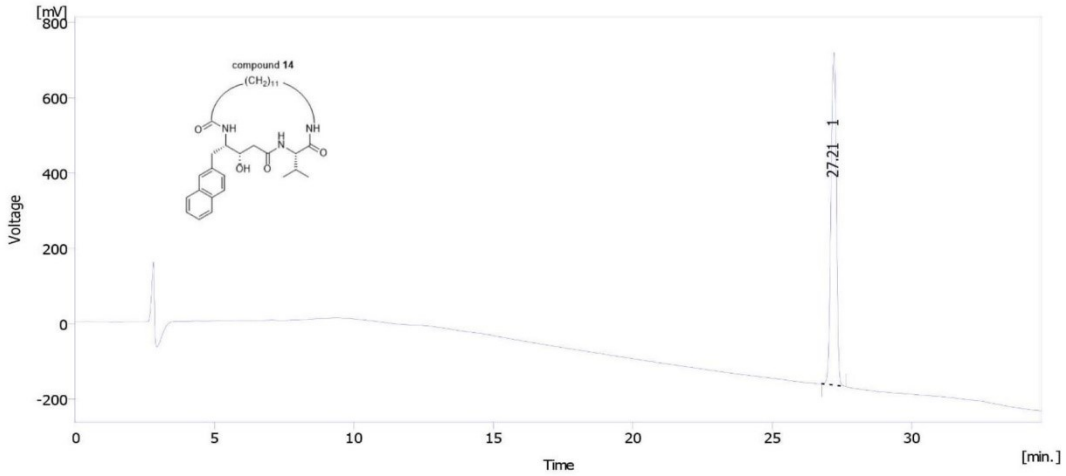
Compound 12



Result Table

	Reten. Time [min]	Area [%]
1	29.584	100.0
	Total	100.0

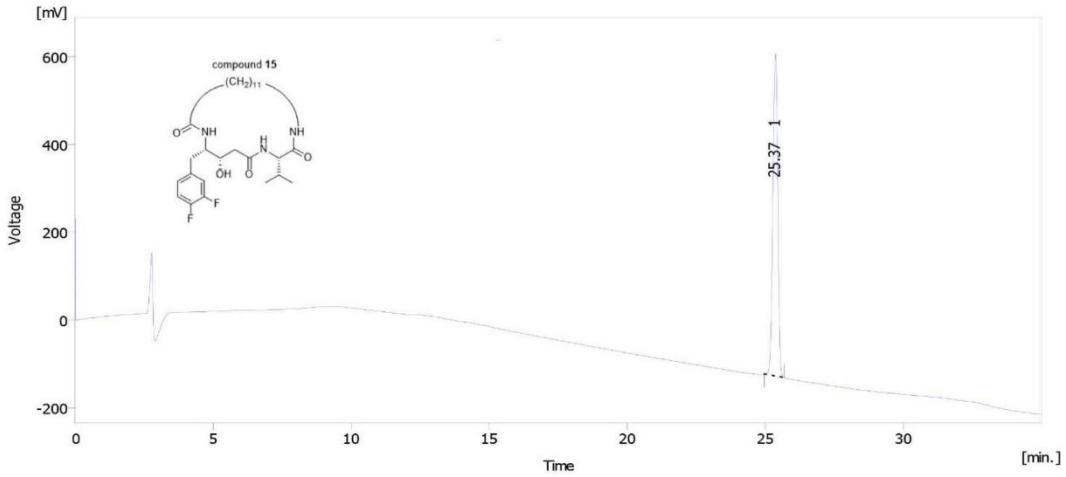
Compound 14



Result Table

	Reten. Time [min]	Area [%]
1	27.212	100.0
	Total	100.0

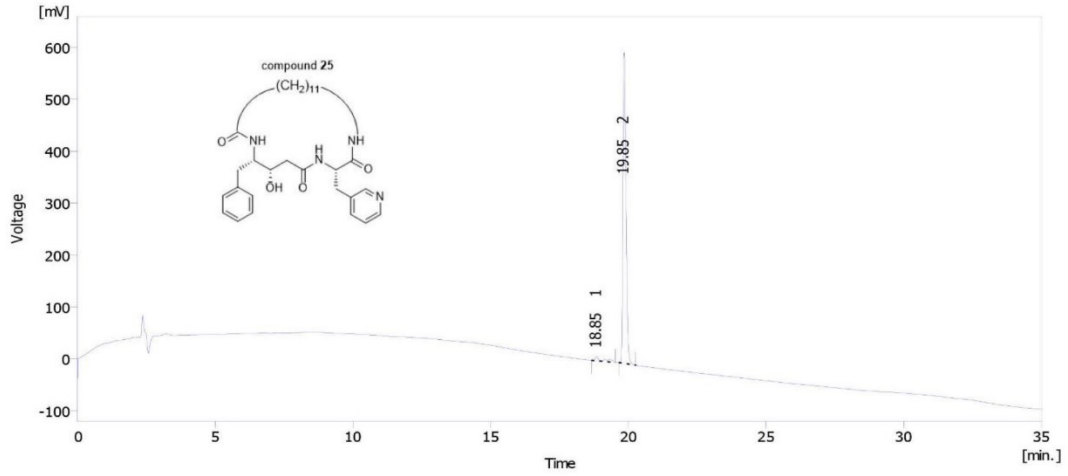
Compound 15



Result Table

	Reten. Time [min]	Area [%]
1	25.372	100.0
	Total	100.0

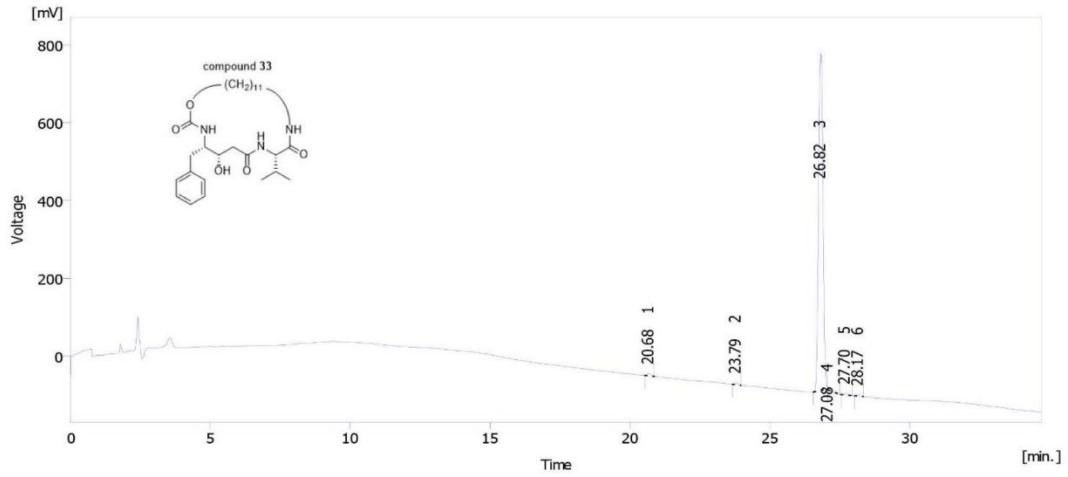
Compound 25



Result Table

	Reten. Time [min]	Area [%]
1	18.852	3.0
2	19.848	97.0
	Total	100.0

Compound 33



Result Table

	Reten. Time [min]	Area [%]
1	20.680	0.5
2	23.792	0.4
3	26.820	98.3
4	27.084	0.2
5	27.696	0.3
6	28.168	0.3
	Total	100.0

5 Diskuze

Katepsin D (KatD) je aspartátová proteasa rodiny pepsinu (rodina A1), jejíž homologa jsou známá u obratlovců i bezobratlých. Lidský KatD se fyziologicky účastní lysozomální degradace proteinů a dějů jako je apoptóza a dále se významně podílí na patofyziologických procesech spojených např. s nádorovými onemocněními nebo Alzheimerovou chorobou[92]. Přestože regulace aktivity této proteasy v buňkách a tkáních je nezbytná, doposud nebyly nalezeny specifické endogenní inhibitory KatD. Známých je pouze několik proteinových a peptidových inhibitorů aspartátových proteas včetně KatD, které ale pocházejí z nižších organismů nebo rostlin a slouží k regulaci exogenních enzymů[214,215]. Předpokládalo se, že kontrola aktivity KatD probíhá především na úrovní ireverzibilní proteolytické aktivace proenzymů na zralý enzym.

V této práci byly identifikovány a popsány první endogenní inhibitory KatD jako přirozené molekuly dvou rozdílných typů, kterými jsou jednak sfingolipidy a jednak peptidy generované z propeptidu KatD. Dále se tato práce zabývala navrhováním biomimetických makrocyklických inhibitorů KatD odvozených z mikrobiálního peptidu pepstatinu.

V publikaci č. 1 jsou sfingolipidy, konkrétně deriváty sfingosinu a ceramidu, charakterizovány jako specifické reversibilní inhibitory lidského KatD. Jejich fosforylované deriváty vykazují naopak aktivační efekt na KatD. V literatuře byl popsán zásadní rozdíl mezi proapoptotickým působením sfingolipidů a antiapoptotickým a mitogenním působením fosfosfingolipidů[216]. Jedná se o (pato)fyziologické procesy, do kterých je zapojen také KatD, a proto je pravděpodobné, že role KatD je v těchto procesech komplexně modulována systémem bioaktivních sfingolipidů a fosfosfingolipidů. KatD je považován za cílovou molekulu pro léčbu nádorových onemocnění (zejména nádoru prsu) a jeho inhibitor ceramid je v literatuře označován jako „tumor supressor lipid“[160,217]. Publikace č. 1 tak přináší zásadní informace pro vývoj nové třídy inhibitorů KatD na bázi sfingolipidů s potenciálním využitím při protinádorové terapii.

Dále je v této práci popsán nový mechanismus modulace aktivity KatD pomocí pH, kdy při interakci s fosfosfingolipidy se zvyšuje celková aktivita KatD a pH optimum enzymu se posouvá k vyšší hodnotě do mírně kyselé oblasti. Tento posun funkčního rozsahu pH by mohl být velmi významný během transportu KatD mimo lysozom za patologických stavů, např. během dobře známé hypersekrece KatD z nádorových buněk do ECM, kde katalyticky aktivní KatD podporuje růst a šíření nádoru[152]. Toto působení je dále

posilováno acidifikací extracelulárního prostředí nádorových buněk v souvislosti se změnami jejich metabolismu[218].

Zatím není znám přesný mechanismus interakce sfingolipidů s KatD. Z výsledků publikace č. 1 plyne, že se jedná o kompetitivní inhibitory, které se vážou alespoň částečně do oblasti aktivního centra, jež obsahuje podmísta preferující vazbu hydrofobních skupin. Přesnou odpověď by poskytlo řešení prostorové struktury komplexu sfingolipidu s KatD. V případě záporně nabitých fosfosfingolipidů lze spekulovat, že interakce s KatD je podobná vazbě záporně nabitých glykosaminoglykanů, které také zvyšují aktivitu KatD[219]. U cysteinových proteas se glykosaminoglykany vážou na kladně nabité části povrchu poblíž aktivního centra a molekula KatD má na svém povrchu také několik takových oblastí[220].

V publikaci č. 2 byl analyzován IrKatD1 z klíštěte a popsán dvojí regulační mechanismus domény propeptidu – intramolekulární a intermolekulární, který využívá vazby propeptidu do tzv. exomísta na povrchu enzymu. Obecně u aspartátových proteas rodiny pepsinu blokuje propeptid přístup substrátu do aktivního centra a intramolekulární autoinhibicí tak kontroluje aktivitu proenzymu. U aspartátových proteas s dlouhým propeptidem (např. proenzym lidského KatD nebo pepsinu) se N-koncová část propeptidu váže do exomísta a C-koncová část propeptidu interaguje pomocí tzv. „lysin-tyrosinové kotvy“ s katalytickými aspartáty[221]. Evoluční extrémní zkrácení propeptidu u proenzymu IrKatD1 vede ke ztrátě přímé interakce s katalytickými aspartáty. Propeptid se stále váže N-koncovou částí do exomísta, ale C-konec pouze stericky brání vstupu substrátu do aktivního centra. Z porovnání prostorových struktur aspartátových proteas plyne, že zatímco C-koncová část propeptidu je u aspartátových proteas rodiny pepsinu vysoce variabilní v sekvenci i délce, N-koncová část propeptidu, která se váže do exomísta, je silně konzervovaná.

Zjednodušená stavba propeptidu IrKatD1 umožnila identifikovat intermolekulární mechanismus inhibice jako nový způsob regulace aktivity proteas rodiny pepsinu. Během procesu autoaktivace proenzymu IrKatD1 dochází k rozštěpení propeptidu na dvě části, přičemž je generován N-koncový oligopeptid, který je pak schopen působit jako silný reversibilní inhibitor zralého rodičovského IrKatD1. Syntetický peptid označovaný [1-8] se sekvencí prvních osmi aminokyselin propeptidu, který nese kritický inhibiční motiv, byl použit k sérii strukturních a funkčních analýz. Z vyřešené krystalové struktury komplexu je vidět, že se peptid [1-8] váže do exomísta na povrchu enzymu. Tím je indukovaná velká konformační změna, která vede k přesunu N-konce zralého enzymu do aktivního centra (na

vzdálenost 35 Å), kde interaguje s katalytickými aspartáty. Peptid [1-8] tak působí jako unikátní alosterický inhibitor zralého KatD1, který funguje odlišným strukturním mechanismem, než je znám z autoinhibice proenzymu IrKatD1.

Schopnost alosterického inhibitoru vázat se na IrKatD1 je závislá na pH: při pH nad 6 vzniká stabilní komplex enzym-inhibitor a s klesajícím pH se jeho stabilita snižuje. Alosterická modulace aktivity IrKatD1 je pravděpodobně významná během biogeneze trávicích vakuol, kdy dochází k výrazným změnám pH[222]. Z proenzymu IrKatD1 je v kyselém prostředí primárních lysozomů (ve střevních buňkách klíštěte) generovaný volný alosterický inhibitor. Následnou fúzí lysozomů s endocytickými vezikuly, které transportují krevní proteiny z lumen střeva, se zvýší pH a alosterický inhibitor se může navázat na IrKatD1 a reversibilně jej inaktivovat. Takto vzniklé trávicí vakuoly se postupně okyselují působením ATPásových pump[222]. V silně kyselém pH se inhibitor uvolňuje z komplexu a katalyticky aktivní IrKatD1 zahajuje štěpení hemoglobinu. Tento scénář je v souladu s předchozími výsledky, které ukazují, že IrKatD1 je klíčový první enzym v celé hemoglobinolytické kaskádě, jejíž průběh a distribuci produktů lze regulovat pomocí přesného nastavení hodnot pH[205].

N-koncová sekvence propeptidu, která rozpoznává exomísto, je vysoce konzervovaná u aspartátových proteas rodiny pepsinu. Peptidy odvozené od této sekvence dokonce křížově inhibují různé proteasy z této rodiny[47]. Tyto výsledky naznačují, že popsaný mechanismus reversibilní regulace aktivity pomocí vlastních alosterických inhibitorů se vyskytuje i u ostatních členů této rodiny. Publikace č. 2 dále definuje strukturní požadavky inhibičního motivu pro alosterickou inhibici a umožňuje tak vývoj nové třídy optimalizovaných inhibitorů cílených na exomísto různých medicinálně významných proteas z rodiny pepsinu.

Publikace č.3 byla zaměřena na racionální navrhování biomimetických inhibitorů lidského KatD, které byly odvozeny z přirozeného mikrobiálního peptidu pepstatinu. Pepstatin jako inhibitor aspartátových proteas je široce používán zejména pro afinitní purifikaci, titraci aktivních center, zobrazovací techniky nebo analýzu inhibiční specifity[205,223–225]. Není ale vhodný jako terapeutické agens kvůli jeho nadměrné velikosti a nepříznivým fyzikálně-chemickým vlastnostem, které způsobují nízkou biodostupnost a rychlou eliminaci[213]. Jde o častý problém lineárních peptidů, který je možné překonat např. konstrukcí cyklických peptidomimetik, jež jsou stále častěji využívány[226].

V této práci byla analyzována změna konformace lineárního peptidu pepstatinu, ke které dochází při jeho interakci s aktivním centrem lidského KatD. Tato konformace byla

stabilizována navržením makrocyklické peptidové kostry, která přírodní vazebný mód mimikuje. Tak byly odvozeny vysoce účinné inhibitory KatD s minimalizovanou velikostí, jejichž způsob interakce byl potvrzen rentgenostrukturní analýzou. Její výsledky jsou v souladu s údaji, že proteasy rozpoznávají peptidové ligandy s β-strukturou a tento typ konformace může být indukován makrocyklizací zlepšující afinitu k proteasám[227].

Připravené makrocyklické inhibitory KatD vykazují příznivé vlastnosti pro vývoj léčiv, které jsou srovnatelné s jinými medicinálně relevantními makrocykly[228,229]. Zejména splňují Lipinského „pravidlo pěti“ a další fyzikálně-chemické parametry látek pro farmakologické využití, kterým původní pepstatin nevyhovuje. Připravené makrocykly byly také úspěšně testovány jako proteolyticky stabilní a netoxické pro buňky a bylo prokázáno, že jejich rozpustnost, buněčná permeabilita a inhibiční specifita se dají účinně modifikovat pomocí cílených substitucí ve struktuře. Získané biomimetické makrocykly tak představují vhodný templát pro další optimalizaci a navrhování potenciálních chemoterapeutik proti patologiím spojeným s lidským KatD a příbuznými proteasami.

6 Závěry

Disertační práce se zabývá novými mechanismy regulace proteolytické aktivity katepsinu D (KatD) a příbuzných aspartátových proteas. Výsledky jsou shrnuty ve třech původních publikacích. Práce přináší zejména objev přirozených endogenních inhibitorů studovaných proteas a dále vývoj syntetických biomimetických inhibitorů. Studované nové typy inhibitorů představují molekulární templáty pro navrhování potenciálních léčiv.

V rámci disertační práce byly splněny zadané cíle s následujícími hlavními závěry:

Regulace aktivity KatD pomocí sfingolipidů (publikace č. 1)

Sfingolipidy byly identifikovány jako první endogenní inhibitory lidského KatD. Detailně byl analyzován vliv struktury sfingolipidů na jejich funkci: sfingolipidy typu sfingosinu a ceramidu účinně inhibovaly KatD, ale naopak jejich fosforylované deriváty byly aktivátory enzymu a ovlivňovaly jeho pH optimum. Komplexní modulace aktivity pomocí sfingolipidů byla specifická pro lidský KatD a u jiných testovaných proteas nebyla zjištěna. Dále byla nalezena korelace mezi působením bioaktivních sfingolipidů na lidský KatD a jejich dříve popsaným modulačním účinkem na nádorové buňky, což naznačuje účast tohoto molekulárního interakčního systému v patofyziologických procesech.

Regulace aktivity KatD pomocí alosterické inhibice (publikace č. 2)

U katepsinu D z klíštěte obecného (IrKatD1) byla pomocí strukturní a biochemické analýzy popsána aktivace neaktivního proenzymu na zralý aktivní enzym. Bylo zjištěno, že autokatalytickou proteolýzou je během tohoto procesu generován peptidový fragment pocházející z propeptidové domény, který funguje jako inhibitor zralého enzymu. Inhibitor se váže do exomísta na povrchu enzymu a působí unikátním alosterickým mechanismem. Jde o první peptidový endogenní inhibitor aspartátových proteas u vyšších organismů. Výsledky naznačují, že popsaná regulace aktivity je evolučně konzervovaná a vyskytuje je i u příbuzných proteas. Jejich exomísto představuje perspektivní molekulární cíl pro vývoj nových alosterických inhibitorů.

Regulace aktivity KatD pomocí biomimetických inhibitorů (publikace č. 3).

Byly navrženy syntetické makrocyklické inhibitory lidského KatD, které napodobují vazebnou konformaci bakteriálního inhibitoru pepstatinu v aktivním místě

aspartátových proteas. U těchto biomimetických inhibitorů byl popsán vazebný mód v aktivním centru lidského KatD pomocí strukturní a výpočetní analýzy. Na základě provedeného rozboru strukturně-funkčních vztahů lze cíleně modifikovat funkční vlastnosti inhibitorů včetně jejich specificity. Připravené makrocykly jsou proteolyticky stabilní a netoxické a splňují kritéria fyzikálně-chemických parametrů pro vývoj chemoterapeutik.

6.1 Prohlášení spoluautorů

Iva Hánová se významně podílela na přípravě publikací uvedených v této disertační práci.

Publikace č. 1: Complex modulation of peptidolytic activity of cathepsin D by sphingolipids.
Žebrakovská I., Máša M., Srp J., Horn M., Vávrová K., Mareš M.
Biochim Biophys Acta, 12, 1097-1104 (2011). (IF = 4,4)

Publikace č. 2: Novel structural mechanism of allosteric regulation of aspartic peptidases via an evolutionarily conserved exosite.
Hánová I., Brynda J., Houštecká R., Alam N., Sojka D., Kopáček P., Marešová L., Vondrášek J., Horn M., Schueller-Furman O., Mareš M.
Cell Chem Biol, 25, 318-329 (2018). (IF = 6,8)

Publikace č. 3: Biomimetic macrocyclic inhibition of human cathepsin D: structure-activity relationship and binding mode analysis.
Houštecká R., Hadzima M., Fanfrlík J., Brynda J., Pallová L., **Hánová I.**, Mertlíková-Kaiserová H., Lepšík M., Horn M., Smrčina M., Majer P., Mareš M.
J Med Chem, 63, 1576-1596 (2020). (IF = 6,1)

……

RNDr. Michael Mareš, CSc.

školitel a spoluautor všech uvedených publikací

7 Seznam použité literatury

1. Clark D, Pazdernik N. Proteomics. In: *Biotechnology*. Elsevier Inc.; 2012:269-304.
2. Rawlings ND. Peptidases. In: Brix K, Stöcker W, eds. *Proteases: Structure and Function*. Springer Vienna; 2013:2-7.
3. MEROPS - the Peptidase Database. Accessed January 18, 2019. https://www.ebi.ac.uk/merops/cgi-bin/family_index?type=P
4. Rawlings ND, Barrett AJ, Bateman A. Asparagine Peptide Lyases. *J Biol Chem*. 2011;286(44):38321-38328.
5. PROSPER: Protease substrate specificity webserver. Accessed January 18, 2019. https://prosper.erc.monash.edu.au/methodology.html
6. MEROPS - the Peptidase Database. Accessed September 16, 2020. https://www.ebi.ac.uk/merops/cgi-bin/statistics_index?type=P
7. Eder J, Hommel U, Cumin F, Martoglio B, Gerhartz B. Aspartic proteases in drug discovery. *Curr Pharm Des*. 2007;13(3):271-285.
8. Rawlings ND, Barrett AJ, Bateman A. MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res*. 2012;40(D1):D343.
9. Tang J, Wong RNS. Evolution in the structure and function of aspartic proteases. *J Cell Biochem*. 1987;33(1):53-63.
10. Rawlings ND, Barrett AJ. Introduction: Aspartic and Glutamic Peptidases and Their Clans. In: Rawlings ND, Salvensen GS, eds. *Handbook of Proteolytic Enzmes*. Third Edit. Elsevier Ltd; 2013:3-19.
11. Goldberg DE. Hemoglobin degradation. *Curr Top Microbiol Immunol*. 2005;295:275-291.
12. Sojka D, Hartmann D, Bartošová-Sojková P, Dvořák J. Parasite Cathepsin D-Like Peptidases and Their Relevance as Therapeutic Targets. *Trends Parasitol*. 2016;32(9):708-723.
13. Simões I, Faro C. Structure and function of plant aspartic proteinases. *Eur J Biochem*. 2004;271(11):2067-2075.
14. Mutlu 1 A, Gal S. Plant aspartic proteinases: enzymes on the way to a function. *Physiol Plant*. 1999;105(3):569-576.
15. Kadek A, Tretyachenko V, Mrazek H, et al. Expression and characterization of plant aspartic protease nepenthesin-1 from Nepenthes gracilis. *Protein Expr Purif*. 2014;95:121-128.
16. Mandujano-González V, Villa-Tanaca L, Anducho-Reyes MA, Mercado-Flores Y. Secreted fungal aspartic proteases: A review. *Rev Iberoam Micol*. 2016;33(2):76-82.
17. Konvalinka J, Kräusslich H-G, Müller B. Retroviral proteases and their roles in virion maturation. *Virology*. 2015;479-480:403-417.
18. Yadati T, Houben T, Bitorina A, Shiri-Sverdlov R. The Ins and Outs of Cathepsins: Physiological Function and Role in Disease Management. *Cells*. 2020;9(7):1679.
19. Gould AB, Skeggs LT, Kahn JR. The presence of renin activity in blood vessel walls. *J Exp Med*. 1964;119(3):389-399.
20. Vassar R. Beta-Secretase Cleavage of Alzheimer's Amyloid Precursor Protein by the Transmembrane Aspartic Protease BACE. *Science (80-)*. 1999;286(5440):735-741.
21. Wisniewski HM, Iqbal K, Bancher C, Miller D, Currie J. Cytoskeletal protein pathology and the formation of beta-amyloid fibers in Alzheimer's disease. *Neurobiol Aging*. 1989;10(5):409-412.

22. Holm I, Ollo R, Panthier JJ, Rougeon F. Evolution of aspartyl proteases by gene duplication: the mouse renin gene is organized in two homologous clusters of four exons. *EMBO J*. 1984;3(3):557-562.
23. Hussain I, Powell D, Howlett DR, et al. Identification of a novel aspartic protease (Asp 2) as β-secretase. *Mol Cell Neurosci*. 1999;14(6):419-427.
24. Törmäkangas K, Hadlington JL, Pimpl P, et al. A Vacuolar Sorting Domain May Also Influence the Way in Which Proteins Leave the Endoplasmic Reticulum. *Plant Cell*. 2001;13(9):2021-2032.
25. Coates L, Tuan H-F, Tomanicek S, et al. The Catalytic Mechanism of an Aspartic Proteinase Explored with Neutron and X-ray Diffraction. *J Am Chem Soc*. 2008;130(23):7235-7237.
26. Aspartic Protease Inhibitors | Cambridge MedChem Consulting. Accessed January 20, 2019. https://www.cambridgemedchemconsulting.com/resources/hit_identification/aspartic_protease_inhibitors.html
27. Nakagawa T, Akaki J, Satou R, et al. The His-Pro-Phe motif of angiotensinogen is a crucial determinant of the substrate specificity of renin. *Biol Chem*. 2007;388(2):237-246.
28. Oka T, Morihara K. Specificity of pepsin: Size and property of the active site. *FEBS Lett*. 1970;10(4):222-224.
29. Dunn BM. Structure and Mechanism of the Pepsin-Like Family of Aspartic Peptidases. *Chem Rev*. 2002;102(12):4431-4458.
30. Metcalf P, Fusek M. Two crystal structures for cathepsin D: The lysosomal targeting signal and active site. *EMBO J*. 1993;12(4):1293-1302.
31. Dunn BM, Hung S. The two sides of enzyme–substrate specificity: lessons from the aspartic proteinases. *Biochim Biophys Acta - Protein Struct Mol Enzymol*. 2000;1477(1-2):231-240.
32. Hofmann T, Hodges RS. A new chromophoric substrate for penicillopepsin and other fungal aspartic proteinases. *Biochem J*. 1982;203(3):603-610.
33. Horimoto Y, Dee DR, Yada RY. Multifunctional aspartic peptidase prosegments. *N Biotechnol*. 2009;25(5):318-324.
34. Khan AR, Khazanovich-Bernstein N, Bergmann EM, James MNG. Structural aspects of activation pathways of aspartic protease zymogens and viral 3C protease precursors. *Proc Natl Acad Sci*. 1999;96(20):10968-10975.
35. Dee DR, Filonowicz S, Horimoto Y, Yada RY. Recombinant prosegment peptide acts as a folding catalyst and inhibitor of native pepsin. *Biochim Biophys Acta - Proteins Proteomics*. 2009;1794(12):1795-1801.
36. Kageyama T, Takahashi K. Rabbit pepsinogens. Purification, characterization, analysis of the conversion process to pepsin and determination of the NH2-terminal amino-acid sequences. *Eur J Biochem*. 1984;141(2):261-269.
37. Foltmann B, Jensen AL. Human Progastricsin. *Eur J Biochem*. 2005;128(1):63-70.
38. Ammerer G, Hunter CP, Rothman JH, Saari GC, Valls LA, Stevens TH. PEP4 gene of Saccharomyces cerevisiae encodes proteinase A, a vacuolar enzyme required for processing of vacuolar precursors. *Mol Cell Biol*. 1986;6(7):2490-2499.
39. Kageyama T, Takahashi K. Activation mechanism of monkey and porcine pepsinogens A. One-step and stepwise activation pathways and their relation to intramolecular and intermolecular reactions. *Eur J Biochem*. 1987;165(3):483-488.
40. Barkholt Pedersen V, Asbaek Christensen K, Foltmann B. Investigations on the Activation of Bovine Prochymosin. *Eur J Biochem*. 1979;94(2):573-580.

41. Schlamowitz M, Peterson LU. Studies on the optimum pH for the action of pepsin on “native” and denatured bovine serum albumin and bovine hemoglobin. *J Biol Chem*. 1959;234:3137-3145.
42. Tanaka T, Yada RY. N-terminal portion acts as an initiator of the inactivation of pepsin at neutral pH. *Protein Eng*. 2001;14(9):669-674.
43. Bohak Z. Purification and characterization of chicken pepsinogen and chicken pepsin. *J Biol Chem*. 1969;244(17):4638-4648.
44. Kunimoto S, Aoyagi T, Morishima H, Takeuchi T, Umezawa H. Mechanism of inhibition of pepsin by pepstatin. *J Antibiot (Tokyo)*. 1972;25(4):251-255.
45. Fujinaga M, Chernaia MM, Mosimann SC, James MNG, Tarasova NI. Crystal structure of human pepsin and its complex with pepstatin. *Protein Sci*. 1995;4(5):960-972.
46. Dash C, Phadtare S, Deshpande V, Rao M. Structural and Mechanistic Insight into the Inhibition of Aspartic Proteases by a Slow-Tight Binding Inhibitor from an Extremophilic Bacillus sp.: Correlation of the Kinetic Parameters with the Inhibitor Induced Conformational Changes. *Biochemistry*. 2001;40(38):11525-11532.
47. Fusek M, Mares M, Vágner J, Voburka Z, Baudys M. Inhibition of aspartic proteinases by propart peptides of human procathepsin D and chicken pepsinogen. *FEBS Lett*. 1991;287(1-2):160-162.
48. Máša M, Marešová L, Vondrášek J, Horn M, Ježek J, Mareš M. Cathepsin D Propeptide: Mechanism and Regulation of Its Interaction with the Catalytic Core. *Biochemistry*. 2006;45(51):15474-15482.
49. Kay J, Gustchina A, Li M, et al. The aspartic proteinase from *Saccharomyces cerevisiae* folds its own inhibitor into a helix. *Nat Struct Biol*. 2000;7(2):113-117.
50. Ng KKS, Petersen JFW, Cherney MM, et al. Structural basis for the inhibition of porcine pepsin by *Ascaris* pepsin inhibitor-3. *Nat Struct Biol*. 2000;7(8):653-657.
51. Mares M, Meloun B, Pavlik M, Kostka V, Baudyg M. Primary structure of cathepsin D inhibitor from potatoes and its structure relationship to soybean trypsin inhibitor family. 1989;25(1):94-98.
52. Volpicella M, Leoni C, Costanza A, De Leo F, Gallerani R, R. Ceci L. Cystatins, Serpins and other Families of Protease Inhibitors in Plants. *Curr Protein Pept Sci*. 2011;12(5):386-398.
53. Guerra Y, Valiente PA, Pons T, Berry C, Rudiño-Piñera E. Structures of a bi-functional Kunitz-type STI family inhibitor of serine and aspartic proteases: Could the aspartic protease inhibition have evolved from a canonical serine protease-binding loop? *J Struct Biol*. 2016;195(2):259-271.
54. Thomas DJ, Richards AD, Kay J. Inhibition of aspartic proteinases by alpha 2-macroglobulin. *Biochem J*. 1989;259(3):905-907.
55. Headey SJ, MacAskill UK, Wright MA, et al. Solution Structure of the Squash Aspartic Acid Proteinase Inhibitor (SQAPI) and Mutational Analysis of Pepsin Inhibition. *J Biol Chem*. 2010;285(35):27019-27025.
56. Christeller JT, Farley PC, Ramsay RJ, Sullivan PA, Laing WA. Purification , characterization and cloning of an aspartic proteinase inhibitor from squash phloem exudate. 67:60-67.
57. Galeša K, Pain R, Jongsma MA, et al. Structural characterization of thyroglobulin type-1 domains of equistatin. *FEBS Lett*. 2003;539(1-3):120-124.
58. Lenarčič B, Turk V. Thyroglobulin Type-1 Domains in Equistatin Inhibit Both and Cathepsin D. *J Biol Chem*. 1999;274(2):563-567.
59. Kratzel M, Schlichtner B, Kirchmayer R, Bernkop-Schnürch A. Simplified Pepstatins: Synthesis and Evaluation of N-Terminally Modified Analogues. *J Med*

Chem. 1999;42(11):2041-2045.
60. Nguyen J-T, Hamada Y, Kimura T, Kiso Y. Design of Potent Aspartic Protease Inhibitors to Treat Various Diseases. *Arch Pharm (Weinheim)*. 2008;341(9):523-535.
61. Gacko M, Minarowska A, Karwowska A, Minarowski Ł. Cathepsin D inhibitors. *Folia Histochem Cytobiol*. 2007;45(4):291-313.
62. Tossi A, Benedetti F, Norbedo S, Skrbec D, Berti F, Romeo D. Small hydroxyethylene-based peptidomimetics inhibiting both HIV-1 and C. albicans aspartic proteases. *Bioorg Med Chem*. 2003;11(22):4719-4727.
63. Tang J. Specific and irreversible inactivation of pepsin by substrate-like epoxides. - PubMed - NCBI. *J Biol Chem*. 1971;246(14):4510-4517.
64. Benedetti F, Berti F, Miertus S, Romeo D, Schillani F, Tossi A. Design, synthesis and preliminary evaluation of peptidomimetic inhibitors of HIV aspartic protease with an epoxyalcohol core. Begtrup M, ed. *Arkivoc*. 2003;2003(14):140-154.
65. PRESS EM, PORTER RR, CEBRA J. The isolation and properties of a proteolytic enzyme, cathepsin D, from bovine spleen. *Biochem J*. 1960;74:501-514.
66. Chen S, Dong H, Yang S, Guo H. Cathepsins in digestive cancers. *Oncotarget*. 2017;8(25):41690-41700.
67. Saftig P, Hetman M, Schmahl W, et al. Mice deficient for the lysosomal proteinase cathepsin D exhibit progressive atrophy of the intestinal mucosa and profound destruction of lymphoid cells. *EMBO J*. 1995;14(15):3599-3608.
68. Koike M, Nakanishi H, Saftig P, et al. Cathepsin D deficiency induces lysosomal storage with ceroid lipofuscin in mouse CNS neurons. *J Neurosci*. 2000;20(18):6898-6906.
69. Guicciardi ME, Leist M, Gores GJ. Lysosomes in cell death. *Oncogene*. 2004;23(16):2881-2890.
70. Follo C, Ozzano M, Mugoni V, Castino R, Santoro M, Isidoro C. Knock-Down of Cathepsin D Affects the Retinal Pigment Epithelium, Impairs Swim-Bladder Ontogenesis and Causes Premature Death in Zebrafish. Escriva H, ed. *PLoS One*. 2011;6(7):e21908.
71. Myllykangas L, Tyynelä J, Page-McCaw A, Rubin GM, Haltia MJ, Feany MB. Cathepsin D-deficient Drosophila recapitulate the key features of neuronal ceroid lipofuscinoses. *Neurobiol Dis*. 2005;19(1-2):194-199.
72. Brindley PJ, Kalinna BH, Wong JY, et al. Proteolysis of human hemoglobin by schistosome cathepsin D. *Mol Biochem Parasitol*. 2001;112(1):103-112.
73. Delcroix M, Sajid M, Caffrey CR, et al. A multienzyme network functions in intestinal protein digestion by a platyhelminth parasite. *J Biol Chem*. 2006;281(51):39316-39329.
74. Morales ME, Rinaldi G, Gobert GN, Kines KJ, Tort JF, Brindley PJ. RNA interference of Schistosoma mansoni cathepsin D, the apical enzyme of the hemoglobin proteolysis cascade. *Mol Biochem Parasitol*. 2008;157(2):160-168.
75. Balczun C, Siemanowski J, Pausch JK, et al. Intestinal aspartate proteases TiCatD and TiCatD2 of the haematophagous bug Triatoma infestans (Reduviidae): sequence characterisation, expression pattern and characterisation of proteolytic activity. *Insect Biochem Mol Biol*. 2012;42(4):240-250.
76. Padilha MHP, Pimentel AC, Ribeiro AF, Terra WR. Sequence and function of lysosomal and digestive cathepsin D-like proteinases of Musca domestica midgut. *Insect Biochem Mol Biol*. 2009;39(11):782-791.
77. Ahn J-E, Zhu-Salzman K. CmCatD, a cathepsin D-like protease has a potential role in insect defense against a phytocystatin. *J Insect Physiol*. 2009;55(8):678-685.

78. Brunelle F, Cloutier C, Michaud D. Colorado potato beetles compensate for tomato cathepsin D inhibitor expressed in transgenic potato. *Arch Insect Biochem Physiol*. 2004;55(3):103-113.
79. Cho WL, Raikhel AS. Cloning of cDNA for mosquito lysosomal aspartic protease. Sequence analysis of an insect lysosomal enzyme similar to cathepsins D and E. *J Biol Chem*. 1992;267(30):21823-21829.
80. Gomes FM, Oliveira DMP, Motta LS, Ramos IB, Miranda KM, Machado EA. Inorganic polyphosphate inhibits an aspartic protease-like activity in the eggs of Rhodnius prolixus (Stahl) and impairs yolk mobilization in vitro. *J Cell Physiol*. 2009;222(3):606-611.
81. Leyria J, Fruttero LL, Nazar M, Canavoso LE. The Role of DmCatD, a Cathepsin D-Like Peptidase, and Acid Phosphatase in the Process of Follicular Atresia in Dipetalogaster maxima (Hemiptera: Reduviidae), a Vector of Chagas' Disease. Rahbé Y, ed. *PLoS One*. 2015;10(6):e0130144.
82. Gui ZZ, Lee KS, Kim BY, et al. Functional role of aspartic proteinase cathepsin D in insect metamorphosis. *BMC Dev Biol*. 2006;6(1):49.
83. Yamamoto K. Cathepsin E and cathepsin D: Biosynthesis, processing and subcellular location. In: *Advances in Experimental Medicine and Biology*. Vol 362. Springer New York LLC; 1995:223-229.
84. Bach A-S, Derocq D, Laurent-Matha V, et al. Nuclear cathepsin D enhances TRPS1 transcriptional repressor function to regulate cell cycle progression and transformation in human breast cancer cells. *Oncotarget*. 2015;6(29):28084-28103.
85. Větvicka V, Vágner J, Baudys M, Tang J, Foundling SI, Fusek M. Human breast milk contains procathepsin D--detection by specific antibodies. *Biochem Mol Biol Int*. 1993;30(5):921-928.
86. Zühlsdorf M, Imort M, Hasilik A, von Figura K. Molecular forms of beta-hexosaminidase and cathepsin D in serum and urine of healthy subjects and patients with elevated activity of lysosomal enzymes. *Biochem J*. 1983;213(3):733-740.
87. Baechle D, Flad T, Cansier A, et al. Cathepsin D Is Present in Human Eccrine Sweat and Involved in the Postsecretory Processing of the Antimicrobial Peptide DCD-1L. *J Biol Chem*. 2006;281(9):5406-5415.
88. Castino R, Delpal S, Bouguyon E, Demoz M, Isidoro C, Ollivier-Bousquet M. Prolactin Promotes the Secretion of Active Cathepsin D at the Basal Side of Rat Mammary Acini. *Endocrinology*. 2008;149(8):4095-4105.
89. Bjelle A, Osterlin S. Cathepsin D activity in bovine articular cartilage, synovial membrane and fluid: degradation of cartilage proteoglycans from same joint. *J Rheumatol*. 1976;3(4):400-408.
90. Poole AR, Hembry RM, Dingle JT. Extracellular localization of cathepsin D in ossifying cartilage. *Calcif Tissue Res*. 1973;12(4):313-321.
91. Hakala JK, Oksjoki R, Laine P, et al. Lysosomal Enzymes Are Released From Cultured Human Macrophages, Hydrolyze LDL In Vitro, and Are Present Extracellularly in Human Atherosclerotic Lesions. *Arterioscler Thromb Vasc Biol*. 2003;23(8):1430-1436.
92. Benes P, Vetvicka V, Fusek M. Cathepsin D—Many functions of one aspartic protease. *Crit Rev Oncol Hematol*. 2008;68(1):12-28.
93. Diment S. Different roles for thiol and aspartyl proteases in antigen presentation of ovalbumin. *J Immunol*. 1990;145(2):417-422.
94. Boya P, Andreau K, Poncet D, et al. Lysosomal Membrane Permeabilization Induces Cell Death in a Mitochondrion-dependent Fashion. *J Exp Med*.

2003;197(10):1323-1334.
95. Johansson AC, Steen H, Öllinger K, Roberg K. Cathepsin D mediates cytochrome c release and caspase activation in human fibroblast apoptosis induced by staurosporine. *Cell Death Differ*. 2003;10(11):1253-1259.
96. Kågedal K, Johansson U, Öllinger K. The lysosomal protease cathepsin D mediates apoptosis induced by oxidative stress. *FASEB J*. 2001;15(9):1592-1594.
97. Nilsson C, Johansson U, Johansson A-C, Kågedal K, Öllinger K. Cytosolic acidification and lysosomal alkalinization during TNF-α induced apoptosis in U937 cells. *Apoptosis*. 2006;11(7):1149-1159.
98. Appelqvist H, Johansson A-C, Linderoth E, et al. Lysosome-mediated apoptosis is associated with cathepsin D-specific processing of bid at Phe24, Trp48, and Phe183. *Ann Clin Lab Sci*. 2012;42(3):231-242.
99. Bidère N, Lorenzo HK, Carmona S, et al. Cathepsin D Triggers Bax Activation, Resulting in Selective Apoptosis-inducing Factor (AIF) Relocation in T Lymphocytes Entering the Early Commitment Phase to Apoptosis. *J Biol Chem*. 2003;278(33):31401-31411.
100. Castino R, Bellio N, Nicotra G, Follo C, Trincheri N, Isidoro C. Cathepsin D–Bax death pathway in oxidative stressed neuroblastoma cells. *Free Radic Biol Med*. 2007;42(9):1305-1316.
101. Haendeler J, Popp R, Goy C, Tischler V, Zeiher AM, Dimmeler S. Cathepsin D and H_2O_2 Stimulate Degradation of Thioredoxin-1. *J Biol Chem*. 2005;280(52):42945-42951.
102. Schestkowa O, Geisel D, Jacob R, Hasilik A. The catalytically inactive precursor of cathepsin D induces apoptosis in human fibroblasts and HeLa cells. *J Cell Biochem*. 2007;101(6):1558-1566.
103. Beaujouin M, Baghdiguian S, Glondu-Lassis M, Berchem G, Liaudet-Coopman E. Overexpression of both catalytically active and -inactive cathepsin D by cancer cells enhances apoptosis-dependent chemo-sensitivity. *Oncogene*. 2006;25(13):1967-1973.
104. Berchem G, Glondu M, Gleizes M, et al. Cathepsin-D affects multiple tumor progression steps in vivo: proliferation, angiogenesis and apoptosis. *Oncogene*. 2002;21(38):5951-5955.
105. Laurent-Matha V, Maruani-Herrmann S, Prébois C, et al. Catalytically inactive human cathepsin D triggers fibroblast invasive growth. *J Cell Biol*. 2005;168(3):489-499.
106. Augereau P, Garcia M, Mattei MG, et al. Cloning and Sequencing of the 52K Cathepsin D Complementary Deoxyribonucleic Acid of MCF7 Breast Cancer Cells and Mapping on Chromosome 11. *Mol Endocrinol*. 1988;2(2):186-192.
107. Cavaillès V, Augereau P, Rochefort H. Cathepsin D gene is controlled by a mixed promoter, and estrogens stimulate only TATA-dependent transcription in breast cancer cells. *Proc Natl Acad Sci U S A*. 1993;90(1):203-207.
108. Westley B, Rochefort H. A secreted glycoprotein induced by estrogen in human breast cancer cell lines. *Cell*. 1980;20(2):353-362.
109. Rochefort H, Cavailles V, Augereau P, et al. Overexpression and hormonal regulation of pro-cathepsin D in mammary and endometrial cancer. *J Steroid Biochem*. 1989;34(1-6):177-182.
110. Elangovan S, Moulton BC. Progesterone and estrogen control of rates of synthesis of uterine cathepsin D. *J Biol Chem*. 1980;255(15):7474-7479.
111. Erickson AH, Conner GE, Blobel G. Biosynthesis of a lysosomal enzyme. Partial structure of two transient and functionally distinct NH2-terminal sequences in

cathepsin D. *J Biol Chem*. 1981;256(21):11224-11231.
112. Nakao Y, Kozutsumi Y, Kawasaki T, Yamashina I, Van Halbeek H, Vliegenthart JFG. Oligosaccharides on cathepsin D from porcine spleen. *Arch Biochem Biophys*. 1984;229(1):43-54.
113. Fortenberry SC, Schorey JS, Chirgwin JM. Role of glycosylation in the expression of human procathepsin D. *J Cell Sci*. 1995;108 (Pt 5:2001-2006.
114. Figura K V, Hasilik A. Lysosomal Enzymes and their Receptors. *Annu Rev Biochem*. 1986;55(1):167-193.
115. Zaidi N, Maurer A, Nieke S, Kalbacher H. Cathepsin D: A cellular roadmap. *Biochem Biophys Res Commun*. 2008;376(1):5-9.
116. Gopalakrishnan MM, Grosch H-W, Locatelli-Hoops S, et al. Purified recombinant human prosaposin forms oligomers that bind procathepsin D and affect its autoactivation. *Biochem J*. 2004;383(3):507-515.
117. Rijnboutt S, Kal AJ, Geuze HJ, Aerts H, Strous GJ. Mannose 6-phosphate-independent targeting of cathepsin D to lysosomes in HepG2 cells. *J Biol Chem*. 1991;266(35):23586-23592.
118. Hiraiwa M, Martin BM, Kishimoto Y, Conner GE, Tsuji S, O'Brien JS. Lysosomal Proteolysis of Prosaposin, the Precursor of Saposins (Sphingolipid Activator Proteins): Its Mechanism and Inhibition by Ganglioside. *Arch Biochem Biophys*. 1997;341(1):17-24.
119. Dittmer F, Ulbrich EJ, Hafner A, et al. Alternative mechanisms for trafficking of lysosomal enzymes in mannose 6-phosphate receptor-deficient mice are cell type-specific. *J Cell Sci*. 1999;112 (Pt 1:1591-1597.
120. Larsen LB, Boisen A, Petersen TE. Procathepsin D cannot autoactivate to cathepsin D at acid pH. *FEBS Lett*. 1993;319(1-2):54-58.
121. Conner GE, Richo G. Isolation and characterization of a stable activation intermediate of the lysosomal aspartyl protease cathepsin D. *Biochemistry*. 1992;31(4):1142-1147.
122. Wittlin S, Rösel J, Hofmann F, Stover DR. Mechanisms and kinetics of procathepsin D activation. *Eur J Biochem*. 1999;265(1):384-393.
123. Richo GR, Conner GE. Structural requirements of procathepsin D activation and maturation. *J Biol Chem*. 1994;269(20):14806-14812.
124. Gieselmann V, Hasilik A, von Figura K. Processing of human cathepsin D in lysosomes in vitro. *J Biol Chem*. 1985;260(5):3215-3220.
125. Samarel AM, Ferguson AG, Decker RS, Lesch M. Effects of cysteine protease inhibitors on rabbit cathepsin D maturation. *Am J Physiol*. 1989;257(6 Pt 1):C1069-79.
126. Laurent-Matha V, Derocq D, Prébois C, Katunuma N, Liaudet-Coopman E. Processing of Human Cathepsin D Is Independent of Its Catalytic Function and Auto-Activation: Involvement of Cathepsins L and B. *J Biochem*. 2006;139(3):363-371.
127. Takahiko K, Koichi H, Shinsei G, et al. Proteolytic processing sites producing the mature form of human cathepsin D. *Int J Biochem*. 1992;24(9):1487-1491.
128. Follo C, Castino R, Nicotra G, Trincheri NF, Isidoro C. Folding, activity and targeting of mutated human cathepsin D that cannot be processed into the double-chain form. *Int J Biochem Cell Biol*. 2007;39(3):638-649.
129. Erickson AH, Blobel G. Carboxyl-terminal proteolytic processing during biosynthesis of the lysosomal enzymes beta-glucuronidase and cathepsin D. *Biochemistry*. 1983;22(22):5201-5205.
130. Horst M, Hasilik A. Expression and maturation of human cathepsin D in baby-

hamster kidney cells. *Biochem J*. 1991;273(Pt 2):355-361.
131. Erickson AH, Blobel G. Early events in the biosynthesis of the lysosomal enzyme cathepsin D. *J Biol Chem*. 1979;254(23):11771-11774.
132. Zaragozá R, Torres L, García C, et al. Nitration of cathepsin D enhances its proteolytic activity during mammary gland remodelling after lactation. *Biochem J*. 2009;419(2):279-288.
133. Reid WA, Valler MJ, Kay J. Immunolocalization of cathepsin D in normal and neoplastic human tissues. *J Clin Pathol*. 1986;39(12):1323-1330.
134. Bossard N, Descotes F, Bremond AG, et al. Keeping data continuous when analyzing the prognostic impact of a tumor marker: an example with cathepsin D in breast cancer. *Breast Cancer Res Treat*. 2003;82(1):47-59.
135. Shin IY, Sung NY, Lee YS, et al. The Expression of Multiple Proteins as Prognostic Factors in Colorectal Cancer: Cathepsin D, p53, COX-2, Epidermal Growth Factor Receptor, C-erbB-2, and Ki-67. *Gut Liver*. 2014;8(1):13-23.
136. Fusek M, Vetvicka V. Mitogenic function of human procathepsin D: the role of the propeptide. *Biochem J*. 1994;303 (Pt 3:775-780.
137. Vetvicka V, Vetvickova J, Fusek M. Role of procathepsin D activation peptide in prostate cancer growth. *Prostate*. 2000;44(1):1-7.
138. Bazzett LB, Watkins CS, Gercel-Taylor C, Taylor DD. Modulation of Proliferation and Chemosensitivity by Procathepsin D and Its Peptides in Ovarian Cancer. *Gynecol Oncol*. 1999;74(2):181-187.
139. Vashishta A, Ohri SS, Proctor M, Fusek M, Vetvicka V. Role of activation peptide of procathepsin D in proliferation and invasion of lung cancer cells. *Anticancer Res*. 26(6B):4163-4170.
140. Achour O, Bridiau N, Kacem M, et al. Cathepsin D activity and selectivity in the acidic conditions of a tumor microenvironment: Utilization in the development of a novel Cathepsin D substrate for simultaneous cancer diagnosis and therapy. *Biochimie*. 2013;95(11):2010-2017.
141. Leto G, Gebbia N, Rausa L, Tumminello FM. Cathepsin D in the malignant progression of neoplastic diseases (review). *Anticancer Res*. 12(1):235-240.
142. Briozzo P, Badet J, Capony F, et al. MCF7 mammary cancer cells respond to bFGF and internalize it following its release from extracellular matrix: a permissive role of cathepsin D. *Exp Cell Res*. 1991;194(2):252-259.
143. Laurent-Matha V, Huesgen PF, Masson O, et al. Proteolysis of cystatin C by cathepsin D in the breast cancer microenvironment. *FASEB J*. 2012;26(12):5172-5181.
144. Laurent-Matha V, Maruani-Herrmann S, Prébois C, et al. Catalytically inactive human cathepsin D triggers fibroblast invasive growth. *J Cell Biol*. 2005;168(3):489-499.
145. Johnson MD, Torri JA, Lippman ME, Dickson RB. The role of cathepsin D in the invasiveness of human breast cancer cells. *Cancer Res*. 1993;53(4):873-877.
146. Laurent-Matha V, Farnoud MR, Lucas A, Rougeot C, Garcia M, Rochefort H. Endocytosis of pro-cathepsin D into breast cancer cells is mostly independent of mannose-6-phosphate receptors. *J Cell Sci*. 1998;111 (Pt 1:2539-2549.
147. Vetvicka V, Vetvickova J, Fusek M. Anti-human procathepsin D activation peptide antibodies inhibit breast cancer development. *Breast Cancer Res Treat*. 1999;57(3):261-269.
148. Vashishta A, Fusek M, Vetvicka V. Possible role of procathepsin D in human cancer. *Folia Microbiol (Praha)*. 2005;50(1):71-76.
149. Glondu M, Coopman P, Laurent-Matha V, Garcia M, Rochefort H, Liaudet-

Coopman E. A mutated cathepsin-D devoid of its catalytic activity stimulates the growth of cancer cells. *Oncogene*. 2001;20(47):6920-6929.
150. Beaujouin M, Prebois C, Derocq D, et al. Pro-cathepsin D interacts with the extracellular domain of the chain of LRP1 and promotes LRP1-dependent fibroblast outgrowth. *J Cell Sci*. 2010;123(19):3336-3346.
151. Rochefort H, Liaudet-Coopman E. Cathepsin D in cancer metastasis: a protease and a ligand. *APMIS*. 1999;107(1):86-95.
152. Maynadier M, Farnoud R, Lamy P-J, Laurent-Matha V, Garcia M, Rochefort H. Cathepsin D stimulates the activities of secreted plasminogen activators in the breast cancer acidic environment. *Int J Oncol*. 2013;43(5):1683-1690.
153. Wolf M, Clark-Lewis I, Buri C, Langen H, Lis M, Mazzucchelli L. Cathepsin D specifically cleaves the chemokines macrophage inflammatory protein-1 alpha, macrophage inflammatory protein-1 beta, and SLC that are expressed in human breast cancer. *Am J Pathol*. 2003;162(4):1183-1190.
154. Briozzo P, Morisset M, Capony F, Rougeot C, Rochefort H. In vitro degradation of extracellular matrix with Mr 52,000 cathepsin D secreted by breast cancer cells. *Cancer Res*. 1988;48(13):3688-3692.
155. Morikawa W, Yamamoto K, Ishikawa S, et al. Angiostatin Generation by Cathepsin D Secreted by Human Prostate Carcinoma Cells. *J Biol Chem*. 2000;275(49):38912-38920.
156. Piwnica D, Touraine P, Struman I, et al. Cathepsin D Processes Human Prolactin into Multiple 16K-Like N-Terminal Fragments: Study of Their Antiangiogenic Properties and Physiological Relevance. *Mol Endocrinol*. 2004;18(10):2522-2542.
157. Ferreras M, Felbor U, Lenhard T, Olsen BR, Delaissé J. Generation and degradation of human endostatin proteins by various proteinases. *FEBS Lett*. 2000;486(3):247-251.
158. Piwnica D, Fernandez I, Binart N, Touraine P, Kelly PA, Goffin V. A New Mechanism for Prolactin Processing into 16K PRL by Secreted Cathepsin D. *Mol Endocrinol*. 2006;20(12):3263-3278.
159. Nicotra G, Castino R, Follo C, Peracchio C, Valente G, Isidoro C. The dilemma: Does tissue expression of cathepsin D reflect tumor malignancy? The question: Does the assay truly mirror cathepsin D mis-function in the tumor? *Cancer Biomarkers*. 2010;7(1):47-64.
160. Dubey V, Luqman S. Cathepsin D as a Promising Target for the Discovery of Novel Anticancer Agents. *Curr Cancer Drug Targets*. 2017;17(5):404-422.
161. Vezenkov LL, Sanchez CA, Bellet V, et al. Structure-Activity Relationships of JMV4463, a Vectorized Cathepsin D Inhibitor with Antiproliferative Properties: The Unique Role of the AMPA-Based Vector. *ChemMedChem*. 2016;11(3):302-308.
162. Li Z, Bao K, Xu H, et al. Design, synthesis, and bioactivities of tasiamide B derivatives as cathepsin D inhibitors. *J Pept Sci*. 2019;25(4):e3154.
163. Vidoni C, Follo C, Savino M, Melone MAB, Isidoro C. The Role of Cathepsin D in the Pathogenesis of Human Neurodegenerative Disorders. *Med Res Rev*. 2016;36(5):845-870.
164. What Happens to the Brain in Alzheimer's Disease? Accessed February 9, 2019. https://www.nia.nih.gov/health/what-happens-brain-alzheimers-disease
165. Selkoe DJ. Alzheimer's Disease: Genes, Proteins, and Therapy. *Physiol Rev*. 2001;81(2):741-766.
166. Gandy S. The role of cerebral amyloid β accumulation in common forms of Alzheimer disease. *J Clin Invest*. 2005;115(5):1121-1129.

167. Sadik G, Kaji H, Takeda K, et al. In vitro processing of amyloid precursor protein by cathepsin D. *Int J Biochem Cell Biol*. 1999;31(11):1327-1337.
168. Schwagerl AL, Mohan PS, Cataldo AM, Vonsattel JP, Kowall NW, Nixon RA. Elevated levels of the endosomal-lysosomal proteinase cathepsin D in cerebrospinal fluid in Alzheimer disease. *J Neurochem*. 1995;64(1):443-446.
169. Cataldo AM, Paskevich PA, Kominami E, Nixon RA. Lysosomal hydrolases of different classes are abnormally distributed in brains of patients with Alzheimer disease. *Proc Natl Acad Sci U S A*. 1991;88(24):10998-11002.
170. Cataldo AM, Nixon RA. Enzymatically active lysosomal proteases are associated with amyloid deposits in Alzheimer brain. *Proc Natl Acad Sci U S A*. 1990;87(10):3861-3865.
171. Paula V de JR de, Guimarães FM, Diniz BS, Forlenza OV. Neurobiological pathways to Alzheimer's disease: Amyloid-beta, TAU protein or both? *Dement Neuropsychol*. 2009;3(3):188-194.
172. Kenessey A, Nacharaju P, Ko LW, Yen SH. Degradation of tau by lysosomal enzyme cathepsin D: implication for Alzheimer neurofibrillary degeneration. *J Neurochem*. 1997;69(5):2026-2038.
173. Bi X, Haque TS, Zhou J, et al. Novel cathepsin D inhibitors block the formation of hyperphosphorylated tau fragments in hippocampus. *J Neurochem*. 2000;74(4):1469-1477.
174. Papassotiropoulos A, Bagli M, Kurz A, et al. A genetic variation of cathepsin D is a major risk factor for Alzheimer's disease. *Ann Neurol*. 2000;47(3):399-403.
175. Ntais C, Polycarpou A, Ioannidis JPA. Meta-analysis of the association of the cathepsin D Ala224Val gene polymorphism with the risk of Alzheimer's disease: a HuGE gene-disease association review. *Am J Epidemiol*. 2004;159(6):527-536.
176. ateroskleróza | Athero. Accessed February 10, 2019. http://www.athero.cz/projekt-medped/pro-pacienty/co-jsou-to-kardiovaskulární-onemocnění/ateroskleróza/
177. Durán MC, Martín-Ventura JL, Mohammed S, et al. Atorvastatin modulates the profile of proteins released by human atherosclerotic plaques. *Eur J Pharmacol*. 2007;562(1-2):119-129.
178. Morgan J, Leake DS. Acidic pH increases the oxidation of LDL by macrophages. *FEBS Lett*. 1993;333(3):275-279.
179. Leake DS. Does an acidic pH explain why low density lipoprotein is oxidised in atherosclerotic lesions? *Atherosclerosis*. 1997;129(2):149-157.
180. Tapper H, Sundler R. Cytosolic pH regulation in mouse macrophages. Proton extrusion by plasma-membrane-localized H(+)-ATPase. *Biochem J*. 1992;281 (Pt 1:245-250.
181. Chen GC, Lau K, Hamilton RL, Kane JP. Differences in local conformation in human apolipoprotein B-100 of plasma low density and very low density lipoproteins as identified by cathepsin D. *J Biol Chem*. 1991;266(19):12581-12587.
182. van der Westhuyzen DR, Gevers W, Coetzee GA. Cathepsin-D-dependent initiation of the hydrolysis by lysosomal enzymes of apoprotein B from low-density lipoproteins. *Eur J Biochem*. 1980;112(1):153-160.
183. Faust PL, Kornfeld S, Chirgwin JM. Cloning and sequence analysis of cDNA for human cathepsin D. *Proc Natl Acad Sci*. 1985;82(15):4910-4914.
184. Baldwin ET, Bhat TN, Gulnik S, et al. Crystal structures of native and inhibited forms of human cathepsin D: implications for lysosomal targeting and drug design. *Proc Natl Acad Sci*. 1993;90(14):6796-6800.
185. Lee AY, Gulnik S V, Erickson JW. Conformational switching in an aspartic proteinase. *Nat Struct Biol*. 1998;5(10):866-871.

186. Grädler U, Czodrowski P, Tsaklakidis C, et al. Structure-based optimization of non-peptidic Cathepsin D inhibitors. *Bioorg Med Chem Lett*. 2014;24(17):4141-4150.
187. Low JD, Bartberger MD, Chen K, et al. Development of 2-aminooxazoline 3-azaxanthene β-amyloid cleaving enzyme (BACE) inhibitors with improved selectivity against Cathepsin D. *Medchemcomm*. 2017;8(6):1196-1206.
188. Imoto T, Okazaki K, Koga H, Yamada H. Specificity of rat liver cathepsin D. *J Biochem*. 1987;101(3):575-580.
189. Dunn BM, Hung S. The two sides of enzyme-substrate specificity: lessons from the aspartic proteinases. *Biochim Biophys Acta*. 2000;1477(1-2):231-240.
190. Scarborough PE, Dunn BM, Guruprasad K, et al. Exploration of subsite binding specificity of human cathepsin D through kinetics and rule-based molecular modeling. *Protein Sci*. 2008;2(2):264-276.
191. van Noort JM, van der Drift ACM. The selectivity of cathepsin D suggests an involvement of the enzyme in the generation of T-cell epitopes. *J Biol Chem*. 1989;264(24):14159-14164.
192. Woessner JF, Shamberger RJ. Purification and properties of cathepsin D from bovine utrus. *J Biol Chem*. 1971;246(7):1951-1960.
193. Kenessey A, Nacharaju P, Ko LW, Yen SH. Degradation of tau by lysosomal enzyme cathepsin D: implication for Alzheimer neurofibrillary degeneration. *J Neurochem*. 1997;69(5):2026-2038.
194. Heinrich M, Neumeyer J, Jakob M, et al. Cathepsin D links TNF-induced acid sphingomyelinase to Bid-mediated caspase-9 and -3 activation. *Cell Death Differ*. 2004;11(5):550-563.
195. BioLib: Biological library. Accessed March 12, 2019. https://www.biolib.cz/cz/taxon/id19328/
196. Volf P, Horák P. Podřád Metastigmata (Ixodida). In: *Paraziti a Jejich Biologie*. Triton; 2007:260-264.
197. File:Ixodid tick structure.jpg - Wikimedia Commons. Accessed March 14, 2019. https://commons.wikimedia.org/wiki/File:Ixodid_tick_structure.jpg
198. ijournal.cz | Klíště – drobný zabiják schovaný ve vysoké trávě. Po přisátí ho ihned otestujte v domácím Borrelia Testu. Accessed March 13, 2019. http://www.ijournal.cz/kliste-drobny-zabijak-schovany-ve-vysoke-trave/
199. Klíště obecné. Accessed March 14, 2019. http://www.atlaszvirat.cz/kliste-obecne-4092
200. Nuttall PA. Wonders of tick saliva. *Ticks Tick Borne Dis*. 2019;10(2):470-481.
201. Electron microscope images - comparing mouthparts. Accessed July 19, 2019. http://www.dartmoorcam.co.uk/dartmoortickwatch/photos/SEM_photos/SEM_photos_CompareMouthparts.htm
202. The Paralysis Tick of Australia - Basic Ixodid Tick Anatomy. Accessed April 1, 2019. http://www.lowchensaustralia.com/pests/paralysis-tick/basic-anatomy.htm
203. Sonenshine D, Anderson J. Mouthparts and digestive system. In: Sonenshine D, Roe M, eds. *Biology of Ticks*. Second. Oxford University Press; 2014:122-162.
204. Sojka D. Krvemlýnek ve střevě klíštěte. *Živa*. 2016;1:10-13.
205. Horn M, Nussbaumerová M, Šanda M, et al. Hemoglobin Digestion in Blood-Feeding Ticks: Mapping a Multipeptidase Pathway by Functional Proteomics. *Chem Biol*. 2009;16(10):1053-1063.
206. Sojka D, Pytelková J, Perner J, et al. Multienzyme degradation of host serum albumin in ticks. *Ticks Tick Borne Dis*. 2016;7(4):604-613.
207. Franta Z, Frantová H, Konvičková J, et al. Dynamics of digestive proteolytic

system during blood feeding of the hard tick Ixodes ricinus. *Parasit Vectors*. 2010;3(1):119.
208. Logullo C, Vaz I da S, Sorgine MH, et al. Isolation of an aspartic proteinase precursor from the egg of a hard tick, Boophilus microplus. *Parasitology*. 1998;116 (Pt 6:525-532.
209. Sorgine MHF, Logullo C, Zingali RB, Paiva-Silva GO, Juliano L, Oliveira PL. A Heme-binding Aspartic Proteinase from the Eggs of the Hard Tick Boophilus microplus. *J Biol Chem*. 2000;275(37):28659-28665.
210. Boldbaatar D, Sikalizyo Sikasunge C, Battsetseg B, Xuan X, Fujisaki K. Molecular cloning and functional characterization of an aspartic protease from the hard tick Haemaphysalis longicornis. *Insect Biochem Mol Biol*. 2006;36(1):25-36.
211. Sojka D, Franta Z, Frantová H, et al. Characterization of Gut-associated Cathepsin D Hemoglobinase from Tick *Ixodes ricinus* (*Ir* CD1). *J Biol Chem*. 2012;287(25):21152-21163.
212. Anantaraju HS, Battu MB, Viswanadha S, Sriram D, Yogeeswari P. Cathepsin D inhibitors as potential therapeutics for breast cancer treatment: Molecular docking and bioevaluation against triple-negative and triple-positive breast cancers. *Mol Divers*. 2016;20(2):521-535.
213. Tumminello FM, Bernacki RJ, Gebbia N, Leto G. Pepstatins: aspartic proteinase inhibitors having potential therapeutic applications. *Med Res Rev*. 1993;13(2):199-208.
214. Mareš M, Fusek M, Kostka V, Baudyš M. Cathepsin D Inhibitor from Potato Tubers (Solanum tuberosum L.). In: *Advances in Experimental Medicine and Biology*. Vol 306. ; 1991:349-353.
215. Moose RE, Clemente JC, Jackson LR, et al. Analysis of Binding Interactions of Pepsin Inhibitor-3 To Mammalian and Malarial Aspartic Proteases. *Biochemistry*. 2007;46(49):14198-14205.
216. Young MM, Kester M, Wang H-G. Sphingolipids: regulators of crosstalk between apoptosis and autophagy. *J Lipid Res*. 2013;54(1):5-19.
217. Li F, Zhang N. Ceramide: Therapeutic Potential in Combination Therapy for Cancer Treatment. *Curr Drug Metab*. 2015;17(1):37-51.
218. Rofstad EK, Mathiesen B, Kindem K, Galappathi K. Acidic extracellular pH promotes experimental metastasis of human melanoma cells in athymic nude mice. *Cancer Res*. 2006;66(13):6699-6707.
219. Beckman M, Freeman C, Parish CR, Small DH. Activation of cathepsin D by glycosaminoglycans. *FEBS J*. 2009;276(24):7343-7352.
220. Horn M, Jílková A, Vondrášek J, Marešová L, Caffrey CR, Mareš M. Mapping the Pro-Peptide of the *Schistosoma mansoni* Cathepsin B1 Drug Target: Modulation of Inhibition by Heparin and Design of Mimetic Inhibitors. *ACS Chem Biol*. 2011;6(6):609-617.
221. James MNG, Sielecki AR. Molecular structure of an aspartic proteinase zymogen, porcine pepsinogen, at 1.8 Å resolution. *Nature*. 1986;319(6048):33-38.
222. Mindell JA. Lysosomal acidification mechanisms. *Annu Rev Physiol*. 2012;74:69-86.
223. Kazakova O V., Orekhovich VN. Isolation of cathepsins D by affinity chromatography (Russian). *Biokhimiya*. 1975;40(5):969-972.
224. Chen CS, Chen WN, Zhou M, Arttamangkul S, Haugland RP. Probing the cathepsin D using a BODIPY FL-pepstatin A: applications in fluorescence polarization and microscopy. *J Biochem Biophys Methods*. 2000;42(3):137-151.
225. Xie D, Gulnik S, Collins L, Gustchina E, Suvorov L, Erickson JW. Dissection of

the pH dependence of inhibitor binding energies for an aspartic protease: Direct measurement of the protonation states of the catalytic aspartic acid residues. *Biochemistry*. 1997;36(51):16166-16172.
226. Tyndall J, Fairlie D. Macrocycles Mimic The Extended Peptide Conformation Recognized By Aspartic, Serine, Cysteine and Metallo Proteases. *Curr Med Chem*. 2012;8(8):893-907.
227. Glenn MP, Pattenden LK, Reid RC, et al. β-strand mimicking macrocyclic amino acids: Templates for protease inhibitors with antiviral activity. *J Med Chem*. 2002;45(2):371-381.
228. Cummings MD, Sekharan S. Structure-Based Macrocycle Design in Small-Molecule Drug Discovery and Simple Metrics To Identify Opportunities for Macrocyclization of Small-Molecule Ligands. *J Med Chem*. 2019;62(15):6843-6853.
229. Giordanetto F, Kihlberg J. Macrocyclic drugs and clinical candidates: What can medicinal chemists learn from their properties? *J Med Chem*. 2014;57(2):278-295.